Turning Points in Construction

Prof. Dr.- Ing./Univ. Tokyo Thomas Bock

Technical University Munich, Chair for Building Realization and Informatics, Faculty of Architecture

Abstract

Socio-technical changes lead to new opportunities. Robots are customizing the industrialized mass production of non/manufacturing industries. Robots ameliorate adverse working and offer better living conditions during demographic changes in societies. Ultimately autonomous systems and robotics could reduce ecological footprint during extreme environmental conditions.

Keywords: automation, autonomy, industrialization, infra free, robotics, socio technical systems

1959: The turning point in industrialization 50 years ago?

The turning point in industrialization was triggered by the introduction of universal automation by industrial robotics. As a prerequisite George Devol patents a playback device for controlling machines, using magnetic recording in 1946. George Devol and Joseph Engelberger design the first programmable robot "arm" and use the term Universal Automation for the first time and later form the world's first robot company. It is called Unimation, Inc. (Universal Automation). Joseph Engelberger and George Devol's robot evolves into the Unimate in 1959. The Unimate robot combines industrial manipulator technology and nascent computer control technology. Unimation Inc. was the first company to make and sell robots to General Motors and General Electric for pick and place operations..

When I met Joseph Engelberger during ISR on 30, November 2005 we discussed the potential of construction robotics. After having listened to my presentation on construction robots he told me that " I can't describe a construction robot, but I know one when I see one". Deploying robotics in construction would enable to customize industrialized construction.

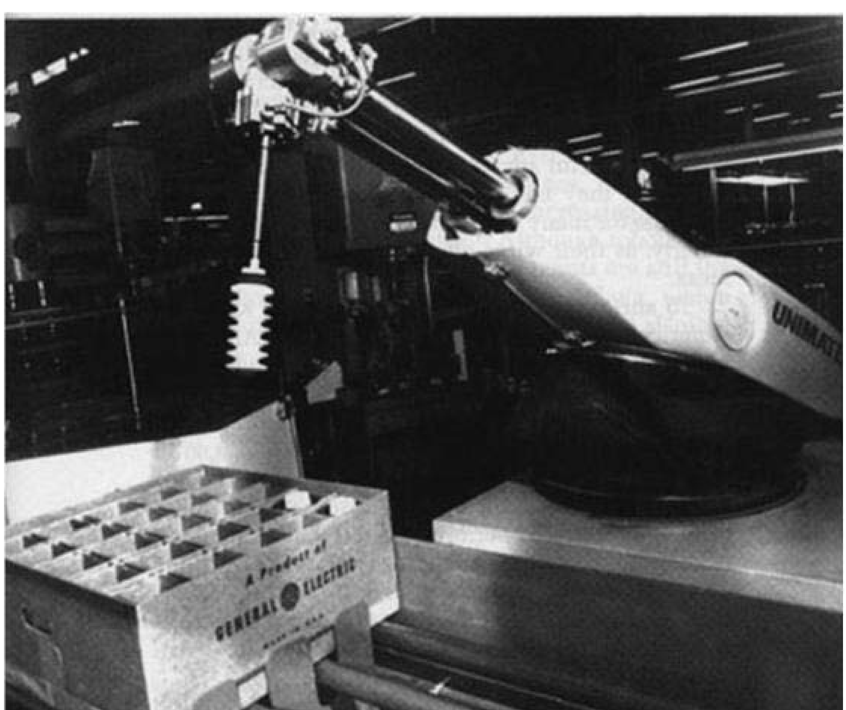


Fig. 01 Unimate robot (copyright J. Engelberger)

1959: The turning point in construction 50 years ago?

Konrad Wachsmann apprenticed as a cabinetmaker and studied at the arts-and-crafts schools of Berlin and Dresden and at the Berlin Academy of Arts. During the late 1920s he was chief architect for Christoph and Unmack, a manufacturer of timber buildings in Niesky. Wachsmann immigrated to the United States in 1941 and went into partnership with the architect Walter Gropius until 1948, an association that resulted in the formation of General Panel Corporation, which produced prefabricated building components.

The philosophy of General Panel System was the same as industrially produced systems such as General Electric, General Motors and JEEP (which originates from GEneral Purpose vehicle). In 1950 he was appointed professor at the Institute of Design of Illinois Institute of Technology, Chicago, and director of the department of advanced building research. His famous written works is The Turning Point of Building (1959; Eng. trans. 1961). In 1964 he joined the USC in L.A. as director of the Building Research Division and chairman of the graduate school of the department of architecture.

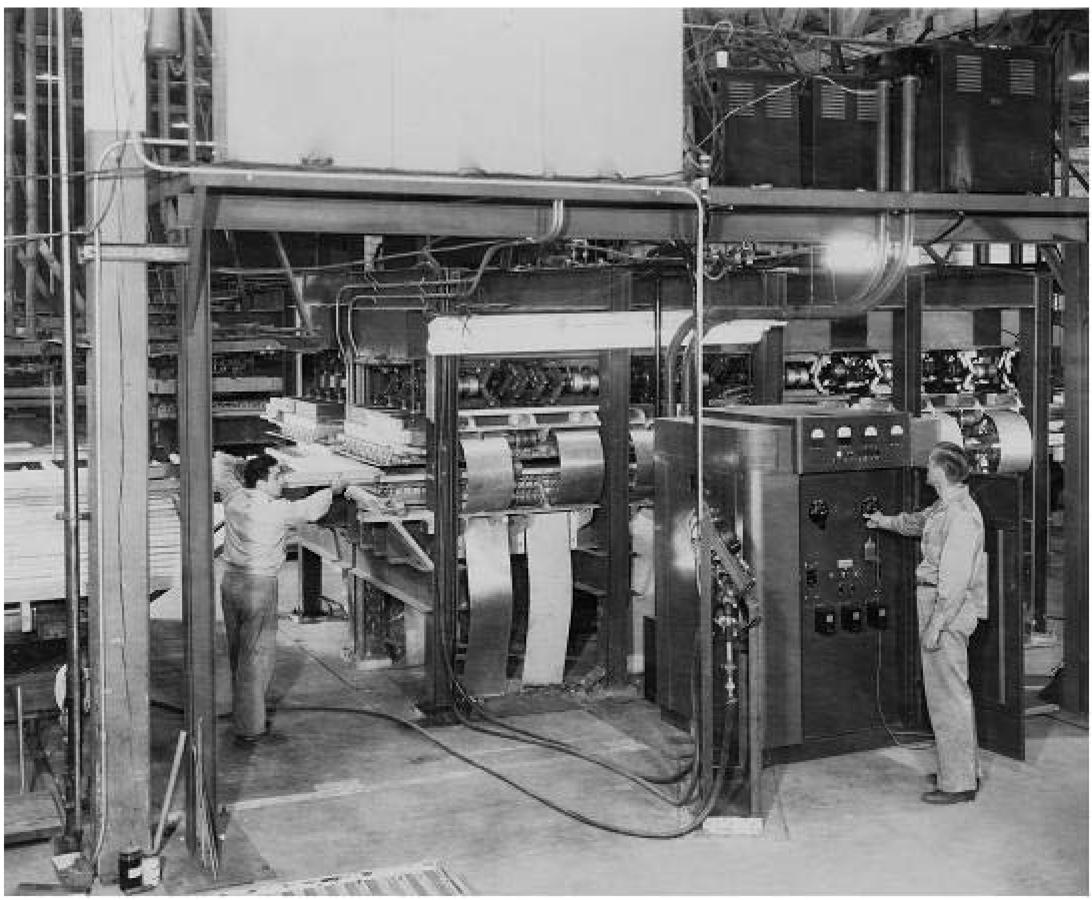


Fig. 02 General Panel Production (Copyright Akademie der Künste Berlin)

1969: The turning point towards robotic Construction 40 years ago?

Konrad Wachsmann's last research project at USC (University of Southern California) in Los Angeles was the Weyerhaeuser sponsored LOM (Location Orientation Manipulator) project between 1969 and 1971. 10 Years after he published his famous book “Turning point of building” he investigated on “motion in time and space”. In October 1989 Fritz Haller, who worked between 1966 and 1970 with Wachsmann, welcomed me at the university Karlsruhe by handing me over a video of the LOM saying:”You go on with it!”The LOM had 7 DOF (degrees of freedom) – as compared to 3-4 DOF of the UNIMATE robot by Joseph Engelberger and George Devol 10 years earlier.

The 7 DOF were realized by 3 orthogonal translatory, 3 Euler angular and 1 polar rotations. Not until recently had there been any 7 DOF industrial robots introduced by Mitsubishi Heavy ind., Yasukawa /

Motoman, DLR, KUKA and Schunck etc. . The 7 DOF kinematic redundancy enables human arm like manipulations resulting in ever broader applications even in ill defined environments.

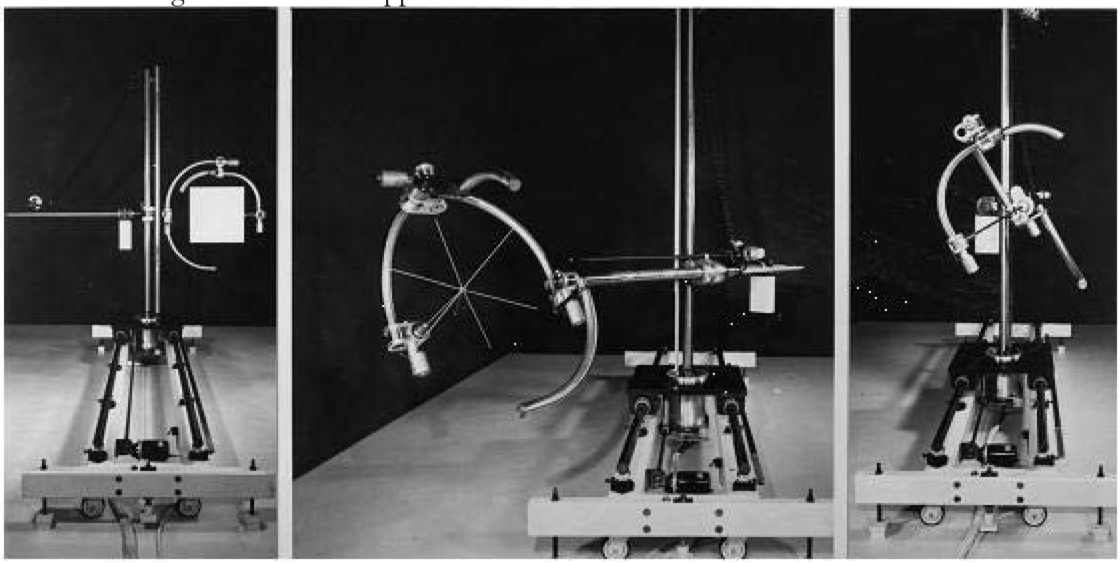


Fig.03 LOM (copyright Akademie der Künste Berlin)

1979: The turning point towards automated construction 30 years ago?

About 30 years ago several groups in Japan began investigating the potential application of construction robotics and over 100 prototypic robots had been build and tested. During my stay in Japan between 1984 and 1989 I analyzed about 50 robots and proposed the notion of robot oriented design of construction products and processes in order to facilitate future construction robotics deployment.

The motivation driving robotization in construction was lack of skilled labor, dangerous-dirty-dull working conditions, no cheap foreign workforce available, high accident and death rates, poor construction quality, time and cost overrun and poor image of construction and building trades in the public opinion.

After a decade of construction robotics development and deployment, the construction industry enjoyed a positive image, social security costs had been reduced, young motivated brightly minded people dreamed of constructing buildings with robots.

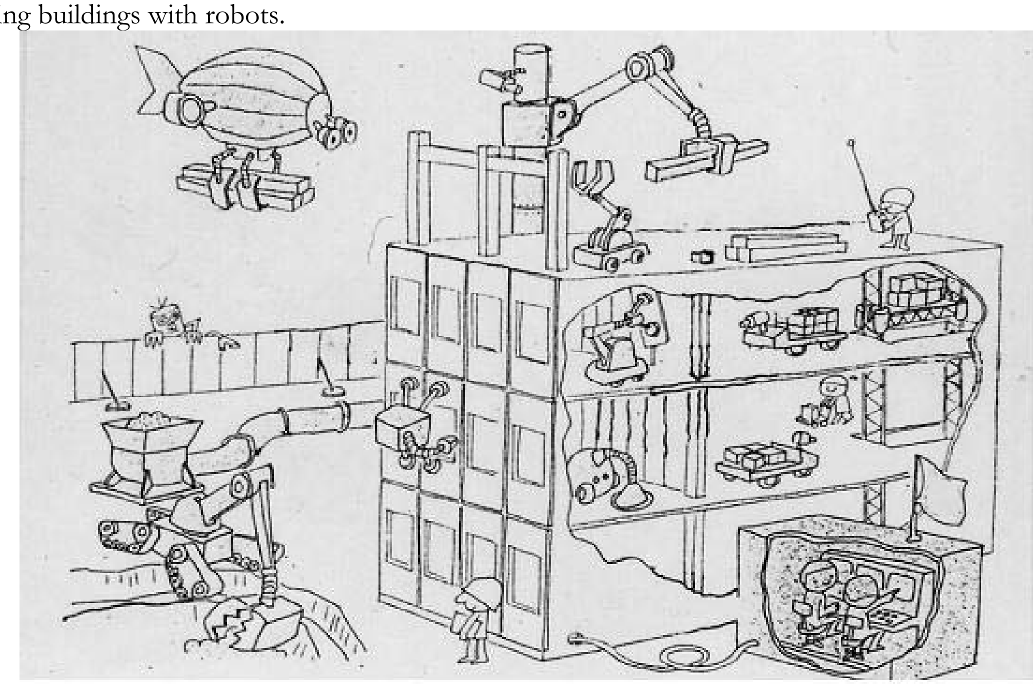


Fig. 04 (AIJ, 1984, Bock)

1989: The turning point towards automated con- and deconstruction 20 years ago?

Three years ago Dr. Paetzold of the “Deutsche Museum” asked me to take a look at a strange mechanism built by the late Konrad Zuse. Having seen some parts I argued that this was an automated building construction system. In 1989 Konrad Zuse, the german computer pioneer, started building the “helixtower” . The helixtower should build itself automatically up and down. First scale models had been realized between 1989 and 1992. Two versions exist and can be seen at the german science museum in Munich. As far as I am concerned Konrad Zuse had in mind assembling and disassembling any kind of building structure automatically.

Being aware of his background as a civil engineer who worked since 1935 for the Henschel aircraft company, I believe that he combined his knowledge of the first computers, which he build since 1936, with the knowledge of lightweight aluminum construction which he probably acquired during his job at Henschel aircraft corporation. Especially the Z 1 -his first computer with its control unit, programming unit and storage devices- in combination with aluminium light weight components, rotational parts magazines and cylindrical up- and downlift mechanism by spiraling rotation converged in his last invention of the helix tower.

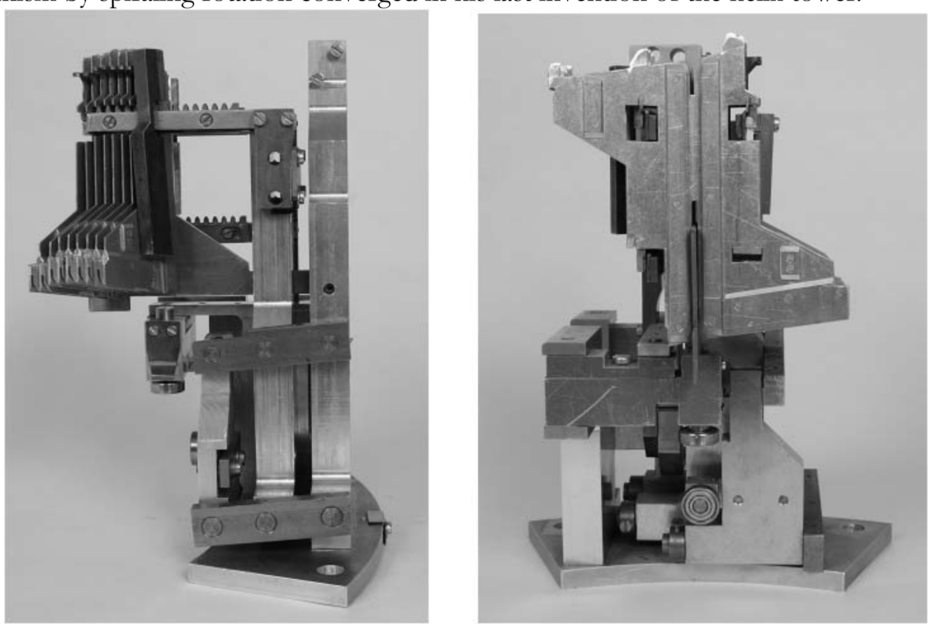


Fig.05 Helixtower (copyright Deutsches Museum);

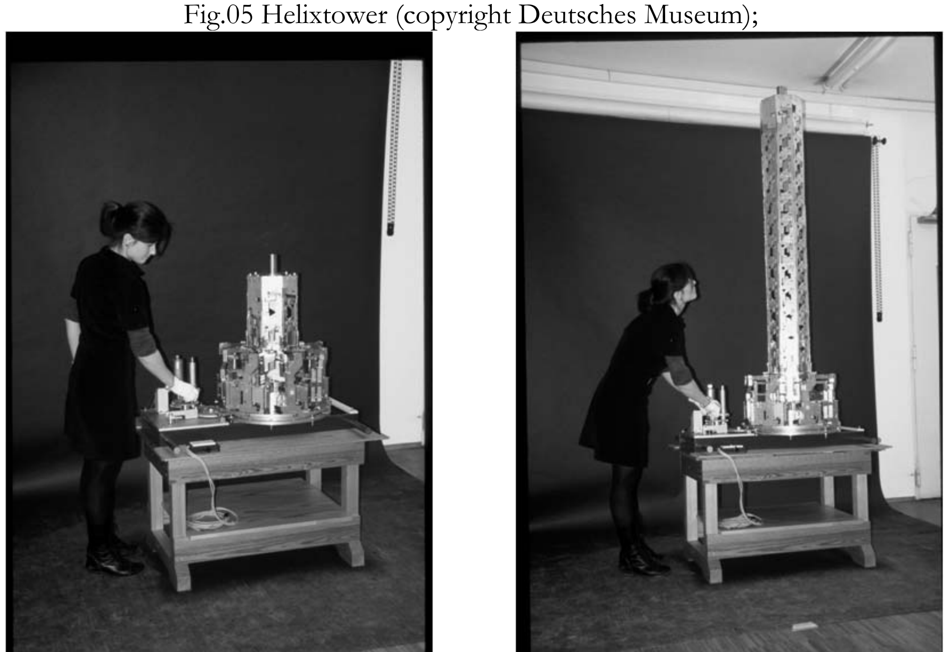


Fig.06 Helixtower, Konrad Zuse (copyright T. Bock)

Helixtower is an automated building construction system similar to Kajima's Amurad automated building construction system of the 1990'ies and the" Daruma otoshi" deconstruction (disassembly-recycling) system of 2008. The helixtower could not just assemble itself but also disassemble itself automatically.

1999: The turning point towards ambient robotics?

In modern times the average life span of a human being has increased substantially mainly due to the higher living standards. This change brings about new challenges for the area of assisted living and the functionality of relevant living space. Current birth rates are not in tune with the requirements of this aging population, leading to a long-term lack of financial provisions as well as lack of appropriate manpower. This demographic revolution requires new socio technical solutions

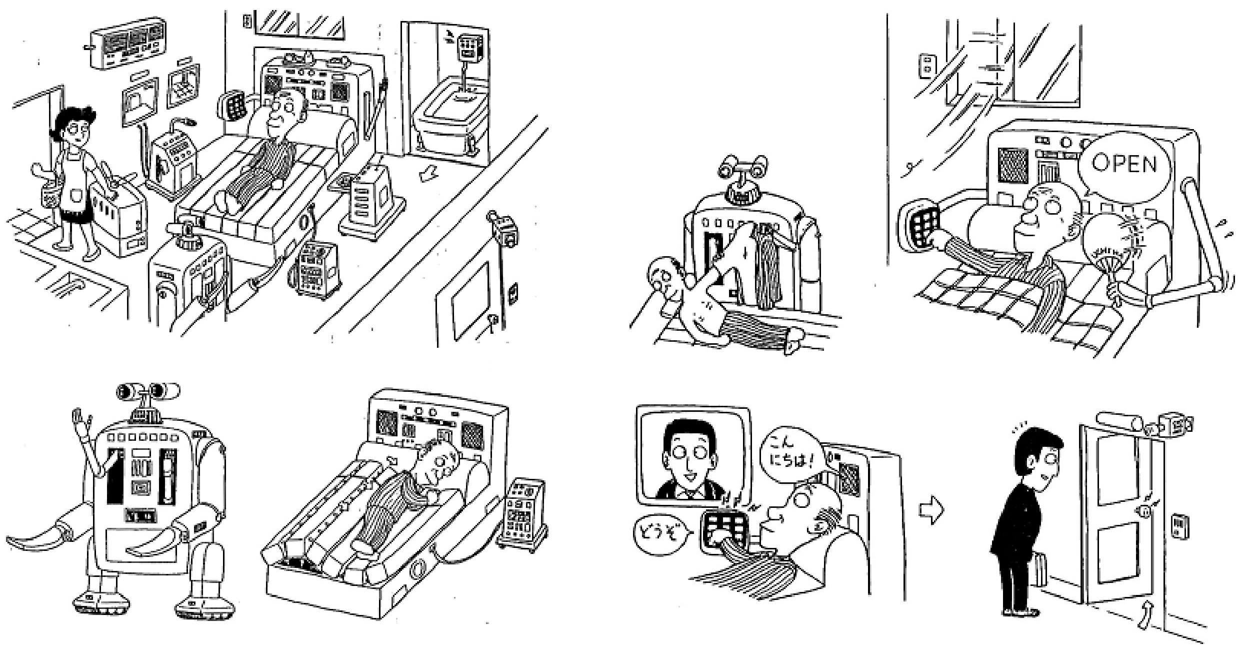


Fig. 07 Sketches showing: Life Support System, by T. Bock, for Japan Science Society1985/8 Sketches: Copyright © by Prof. Dr.-Ing./Univ.Tokio T. Bock

There is a strong tendency of the elderly to stay in their familiar habitat. Technological advancement in the assisted living industry is currently vastly outpacing architectural effort to create effective harmonised living environments. It is important to include these long-term requirements at an early stage of real estate development to overcome higher costs at later stages and possible lack of functionality.

Frontier scientists, engineers, sociologists, architects, robotic specialists, people from the areas ICT, mechatronics and microelectronics, service scientists and even people from the field of medicine, psychology and sociology have to join collaborative and highly interdisciplinary workgroups to develop acceptable ambient robotic environments.

2009: Turning point towards urban robotics

Above mentioned demographic changes in major industrialized countries the ambient robotics development will further permeate into urban environment by robotic cars and human-robot-car-city communication. A sensor network will support intelligent robots servicing in daily human life environment. One sensor network is being developed in the framework of the robot town project. The goal of robot town is to enable robots to execute various tasks for ordinary human life by implementing human tracking by vision systems, distributed sensors and RFID tags thus creating an environment well structured in informative way.

Fig.08 Robot Town (copyright Tsutomo Hasegawa)

2019: Turning point towards space robotics

The planning and building in the aerospace questions the methods of planning and building on earth. To learn anything about space-related building means to think in closed systems. A space station works completely self-sustaining.

The methods of prefabrication, of compact transporting, of modular building and of modification could give an inspiration for the methods of building on earth. Every transaction of execution has to be planned very carefully because nobody could quickly return to earth just to get any forgotten material or tools. At each terrestrial construction site much time is wasted for bringing and taking tools and material.

As space-related building is very expensive, nobody can waste space easily. Because of the zero gravity space can be used ideally in three dimensions, not only on the base side. Every room is multifunctional; the living room is working place, bedroom and dining room at the same time. In this point of view we can question our room concepts on earth and reduce our environmental waste of land. So if we are able to plan, to build and to run a space station, we can also build more ecologically.

2029: Turning point towards infra free life

Long- run missions in the aerospace call for a new idea of supply. The production of nourishment as well as water and air conditioning are tasks which massively influence the design of future space stations. It has to get by with the material and food which it can carry along. Because of the transporting system (load capacity of a skyrocket is about five tons) and the resulting restrictions contemporary space stations are formed by a high rate of prefabrication and an appropriate modular way of light weight construction.

Beside the technical solution biological answers have been considered for over 50 years by Russian scientists (BIOS-3), the Americans by building the biosphere 2 in the desert of Arizona and Japanese authority by biosphere J. Biological processes will shape the architecture of space stations in the future. Autonomous, self reliant and sustaining buildings will avoid energy loss due to infra free entities that generate and recycle habitats own ressources. A long-run mission of space ship earth becomes only imaginable, if technical products get more similar to life sciences or if we utilize organic products in a self sustaining way.

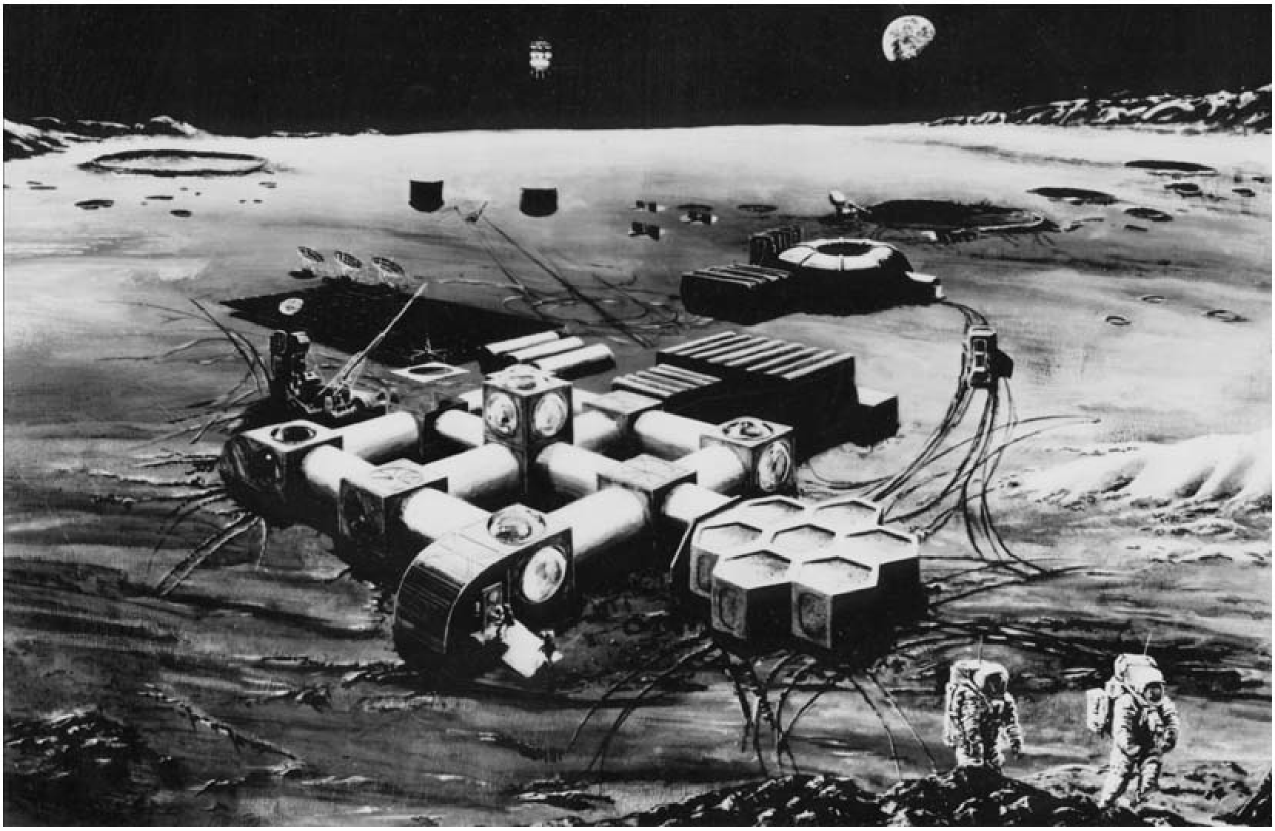


Fig.09 Space Colony (copyright Shimizu)

Literature

[1] S.Anilir, T.Bock, T.Linner, H.Suay, Application of Infra-Free Motherboard (IFM) in a Decentralized Community for a Transformable and Real-time processing Multidirectional Energy Supply Network, May 2009

[2] T. Bock: Das Dach wird zuerst gebaut und das Erdgeschoß zuerst rückgebaut, in: Volume 84, Bauingenieur, Springer ed., Feb. 2009, ISSN 0005-6650

[3] T. Bock; T. Linner: Service Oriented Architecture, 2nd AAL Symposium VDE, Berlin, 2009

[4] Nora Eibisch: Helixturm von Konrad Zuse, Diploma Thesis, TU Munich, 2009

[5] T. Hasegawa, K. Murakami: Robot Town Project, 3rd URAI 2006

[6] Bock, Thomas; Matsumura, Shuichi: Rojin Oriented Design for Architectural Life Support System. - In: Proceedings POLIS/BAS Conference "Universal Design for Buildings: Tools and Policy", Brugge/Gits 16./17.11.06

[7] T. Bock: Robotic Gerontotechnology, Research Project for Japan Science Society, TU Munich, Germany, 2000

[8] Yukio Hasegawa: Robotization in Construction, 10.June 1999, ISBN 4-7693-2145-7

[9] Life Support Sytem (LSS 3), Research report for Japan Science Society, Prof. Thomas Bock, University of Karlsruhe, Germany, 1992

[10] Otto Maier: Die räumliche Syntax, Dissertation , Universität Karlsruhe, 1989

[11] Bock, ThomasL'enjeu des robots et des automatismes- Avances pour le genie urbain, in: Ecole nationale des ponts et chaussees, Laboratoire techniques, territoires et societies unite associee au CNRS, UA 1245, La courtine, BP 105, 93194 Noisy le grand, Cedex, (CRIIF, UPMC , Jussieu, Paris), (Franzšsisch)

[12] T. Bock: Robot Oriented Design, SEKO, Shokokusha Publishing, Tokyo, May, 1988

Managing Data from Instrumentation in the CEDEX Test Track

Javier Perez Ayuso [1], Belen Cadavid Jauregui [2] and Angel Mateos [3]

[1] Transport Research Center of CEDEX, Autovía de Colmenar Viejo, km 18.2, El Goloso, 28760 Madrid, Spain, Email: Javier.Perez@cedex.es
[2] Transport Research Center of CEDEX, Autovía de Colmenar Viejo, km 18.2, El Goloso, 28760 Madrid, Spain, Email: Belen.Cadavid@cedex.es
[3] Transport Research Center of CEDEX, Autovía de Colmenar Viejo, km 18.2, El Goloso, 28760 Madrid, Spain, Email: amateos@cedex.es

Abstract

One of the main objectives of the Test Tracks is to determine the structural response of the pavements under particular load and environmental conditions. Structural response can be measured in terms of different variables, like deflection, stress or strain at certain significant points of the pavement structure. Other related environmental variables as temperature or moisture condition are typically measured as well.

A minimum number of sensors are required for any variable to be measured, since the variability of the response can be significant. Consequently, the number of sensors used in a pavement structural response assessment can be very high.

Response is typically determined for different conditions of load, speed, offset, temperature, number of cycles etc. This requires a very comprehensive planning for the measurements.

This paper presents the Software Program used in CEDEX Test Track to measure and manage the data from the six pavement sections of the facility. The Program enables the planning of measurements for specified conditions of speed, offset, pavement temperature, date or number of cycles, and the later treatment of the registered signals.

The Program can manage up to 256 sensors, and typically around 250.000 measurements are carried out in a complete test. This way, a detailed knowledge of pavement behavior under a wide variety of environmental or load conditions can be achieved.

Introduction

Accelerated Pavement Testing can be defined as the “controlled application of wheel loading to pavement structures for the purpose of simulating the effects of long-term in-service loading conditions in a compressed time period”. There are twelve full-scale facilities operating in Europe and a similar number in the United States of America, besides other facilities in Mexico, Brazil, South Africa, Australia, New Zeeland, China and Japan. It can be stated that nowadays the Accelerated Pavement Testing constitutes a basic pillar of road research worldwide. There are two kinds of test facilities: circular and linear shaped.

Figure 1. Circular Test Track (LCPC France)

Figure 2. Linear Test Track (LINTRACK, Delf, Netherlands)

The CEDEX Test Track is in between, having two straight sections of 75m each, joined by two additional curved sections with a radius of 25 m. A rail beam located on the inside perimeter of the track serves as a guide for two automatic vehicles.

Considering that the six sections of pavements under test are installed on the straight segments of the track, this facility could be classified from the test point of view inside the second group of linear-shaped facilities. The total length traveled by the load test wheel is of 304 meters by cycle. The curved segments are not used for test purposes and are assigned to other studies like surface materials, surface treatments, paints, wearing paths, etc.

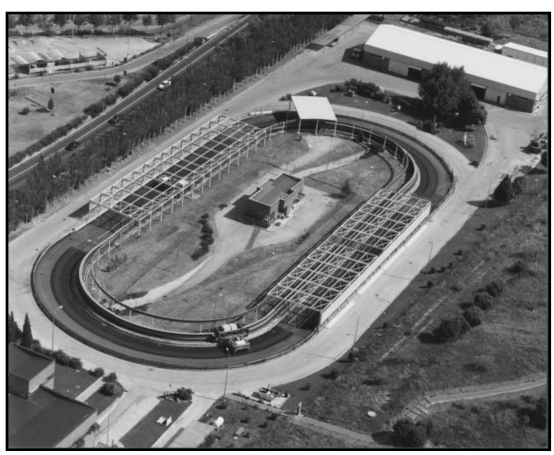


Figure 3. CEDEX Test Track

The testing of the pavement sections is carried out in the straight stretches, and therefore the results are comparable to those obtained in other linear test tracks. Six 20-25 m long complete pavement sections can be tested simultaneously.

Meanwhile the curved segments are based on the terrain; the straight segments are installed inside two watertight U-shaped test pits made out of reinforced concrete. The concrete test pit, 2.6 m deep and 8 m wide, enable the building of embankments of at least 1.25 m in height as well as the use of conventional machinery and the usual road-building procedures. The purpose of using concrete test pits is to isolate the performance of the pavements from that of the surrounding ground, allowing homogeneous support to the pavements throughout each test and between different tests in such a way that the results are comparable. It also allows the subgrade to be flooded for testing under different groundwater conditions.

The two test vehicles apply the load by gravity through a half heavy axle. The load can be fixed in between 5.5 and 7.5 Tons and is fixed at 6.5 Tons, equivalent to the maximum allowed load in Spain for a single axis (13 Tons). The suspension system is pneumatic. The test wheel is equipped with two twin wheels or one single balloon-shaped wheel at 8,5 Kg/cm2 inflating pressure. Both the suspension system so as the traction test wheels are conventional ones used in Roads' Transport. The circulating speed is of 40 Km/h, with a maximum allowable speed of 60 Km/h.

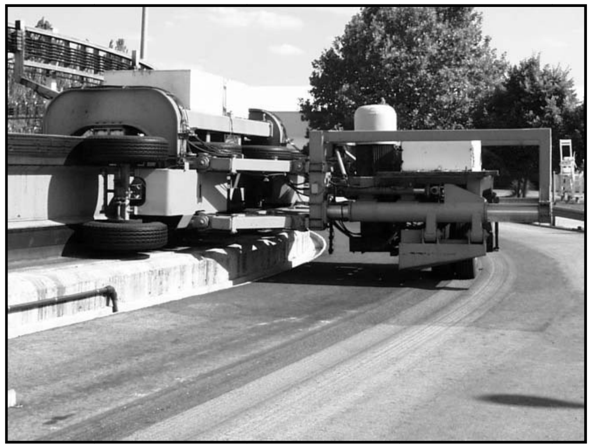


Figure 4. Vehicle for Traffic Simulation

Thanks to a hydraulic actuator, the test wheel can be positioned in seven different transversal paths, producing in this way a footprint band of 1.0 to 1.3 m width. The automatic control of this position allows reproducing a real statistical distribution of passages according to real traffic.

This test facility is fully controlled from a centralized Control Center, situated in the geometrical center of the test track. The control program has been specifically developed for this application. In this way the whole facility can work unattended once programmed. The test frequency is higher than 1x106 cycles by year. The automation process was mainly considered during the design phase of the facility with the aim of reaching a test life cycle with minimum interruptions.

Measuring Parameters

When a wheel moves along a road, stresses and strains develop at any point of the pavement structure; this stresses and strains depend on the type, magnitude and direction of the load, pavement structure, type of subgrade, temperature, depth, etc.

The instrumentation of the pavement makes possible the measurement of the stresses and the strains that appear in different parts of the pavement under the pass of a load, and especially those that are considered to be critical.

For each layer, the critical points as well as the tensodeformational variables are different, and that has to be considered when choosing the type of sensor and its placement.

Horizontal tensile strain at the bottom of the bituminous layer is considered the most important response variable for flexible pavements. Consequently, the instrumentation of the asphalt mixture layers is mainly focused on measuring horizontal strain at the bottom of the layer.

Granular layers and soils fail mainly due to accumulation of vertical strains. Therefore, the instrumentation of the soils is especially focused on measuring vertical stresses and strains.

Pavement deflection sensors are also placed in order to measure the transient response under the pass of the moving wheel. These sensors are placed on the top of the asphalt layer and anchored to the bottom of the test pit.

Finally, a series of sensors are installed in order to collect data from environmental and load related variables: temperature, moisture and water table, speed, transverse position, etc.

Description of the Control System of the Facility

In the Test Track facility, there are two control systems related to one another and both of them are managed by a single system (PC Computer). The first one (PLC Computer) is in charge of the steering of vehicles and it controls the following key test parameters: speed, load transverse position, air pressure of the tires, etc., and also all the required variables for the maintenance of the vehicles and safety of the facility, including: electricity consumption, puncture detectors, position detectors, and so forth. The other system (Micro Computer) manages the instrumentation covering the tasks mentioned above.

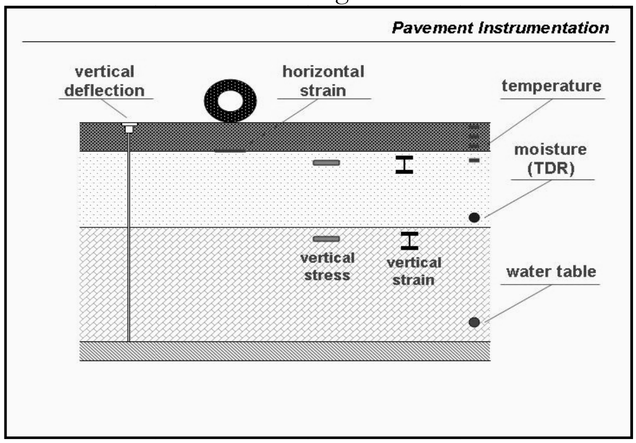


Figure 5. Example of the instrumentation

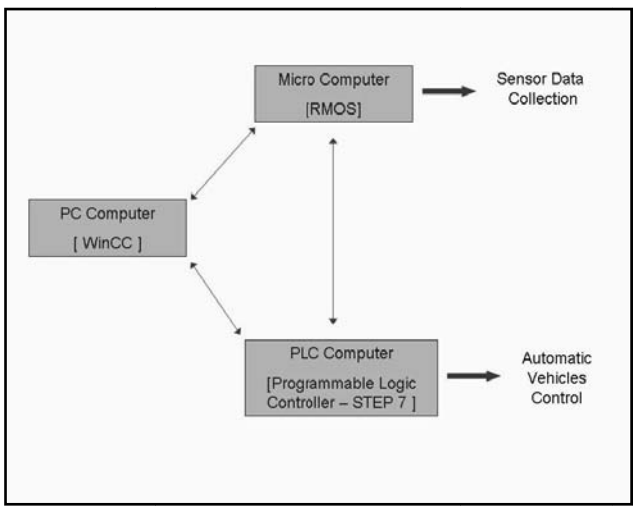


Figure 6. Communication system

Both systems are in permanent communication. The system that controls the instrumentation needs at all times the data regarding the position of the vehicles on the track, the speed and the transverse position. On the other hand, the vehicles control system has to obey the orders issued by the measurement system to achieve the vehicle speed and transverse position required by the measurement.

In the Test Track, a gallery is available on the straight stretches next to the test sections where the power supply and conditioning devices of the sensors are located. It was thus designed to achieve the minimum distance possible between the sensor and the conditioning device to avoid distortion of the output signal and to amplify this signal to reach the Control Center, where the signal will be converted from analog to digital.

Instrumentation Management Tasks

The system for data acquisition of the sensors, which is fully automated, has been designed and developed by CEDEX and makes real-time measurement and storage in database possible for up to 300 sensors on every measurement test.

The management tasks are divided into three processes:

- Management sensor process.
- Measurement process.
- Data storage and analysis.

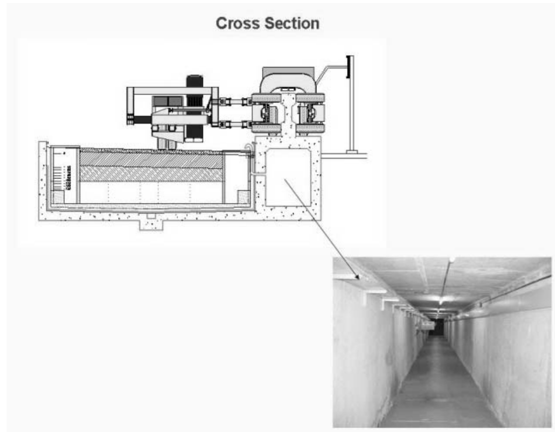


Figure 7. Cross section and instrumentation gallery

Management Sensor Process

Once the plan of instrumentation for each test has been designed and the supply of each sensor has been carried out, the next step is to register it into the system database. It is then time to include all the data which define the sensor, its location, the plate and channel of the measurement chain where the sensor is connected, the optical measurement startup sensor, calibration data of the sensor, relevant dates and state of activity.

After the sensor has been registered in the database, the next step is carrying out the calibration. During the calibration process, the transducer must be calibrated and also the whole measurement chain, including wiring, the conditioning plate and the ADC board. Calibration is performed from the system itself and the results are incorporated into the system database.

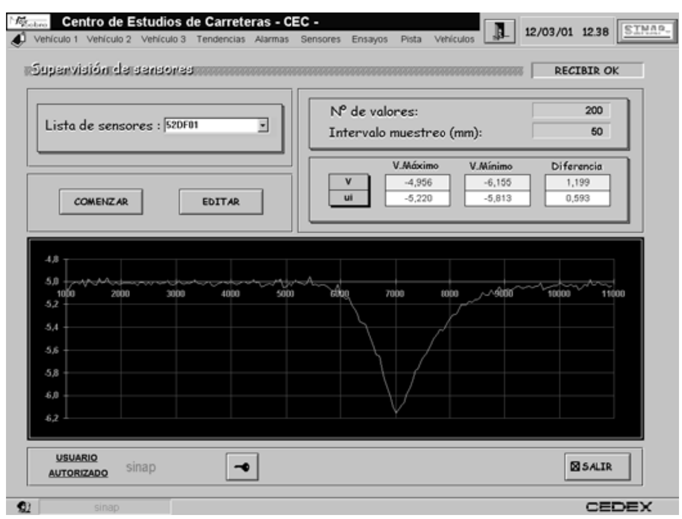


Figure 8. Example of sensor graphical monitoring

There are also additional tasks which simplify the instrumentation maintenance work and feature graphical and numerical real-time monitoring functions of the readings.

Measurement Process

Within the data acquisition system, there are two distinct types of tests:

- Dynamic tests.
- Special tests.

The dynamic test refers to the systematic measures taken with the instrumentation while the vehicle is in motion. When carrying out this kind of tests, the sensors being measured have to be previously defined, up to a number of 256 per test. Besides, a results file (ASCII) is created to include all the variables required for the analysis of curves, such as pavement temperature, ambient temperature, number of cycles, transverse position, vehicle speed, date and time.

Dynamic tests are activated automatically. They can be activated by means of three different events, which are selected when the test is being scheduled. These three events are the following:

- Number of cycles. The test begins when the vehicles cover a predetermined number of cycles.
- Time. The date and time of the beginning of the measurement are indicated.
- Temperature. It begins when the pavement temperature (defined by the user) reaches a certain value.

The presence of staff at the facility is not required for this kind of tests, which are performed 24/7.

When the event that triggers a test takes place, the computer in charge of the measurement management (Micro computer) instructs the computer in charge of the steering of vehicles (PLC computer) to position the vehicles on the conditions required for the test. Once the vehicles are placed in the right position for the scheduled test, an OK signal is sent to the Micro computer, which activates the beginning of the data collection.

The vehicles will complete one cycle, in which the vehicle speed is measured and the sampling speed is calculated in order to storage 5 cm-distance readings which begin 5 m before the vehicle arrives to the vertical position of each sensor and finish when the vehicle moves 5 m onwards the location of the sensor. In the subsequent cycle, a measurement per sensor is carried out, which stores 200 values that define the curve. Once every piece of data is stored, the test finishes and the computer in charge of the steering of the vehicles regains control. This test can be scheduled cyclically depending on the number of cycles, after a determined period of time or when the desired temperature values occur.

We would like to highlight the following three types of tests as special tests:

- Temperature traces. The test starts and finishes automatically at a specified time and is repeated after an elapse of time. In this test, the daily temperature traces are measured for each of the sensors located on the Test Track, which are used to analyze not only the instrumentation but also the damages of the pavement.
- Manual Start. These tests can be performed not only with vehicles in motion but also with vehicles stopped. They begin when a trigger is sent. This kind of test is used to study in detail some specific variables when the vehicle is passing by and, unlike the dynamic tests; they are made with a sample frequency up to 5000 samples per second. This test is also used to measure the response of one or various sensors to equipment other than the test vehicles, e.g. FWD devices.
- Start by optical sensor. It has the same features as the Manual Start test, but, in this case, the measurement is triggered by one of the optical sensors on the Test Track.

Data Storage and Analysis

All the files generated by the measurement process are included in an ORACLE database for subsequent analysis. The only measurements that are stored come from the systematized dynamic tests, thus the special tests results are kept for the user personalized processing. At this stage, the system has the three following distinct processes:

- Data entry
- Curve display
- Parameter analysis

Data entry

The data measured during the measurement process are not always correct and are subject to wrong readings caused by the failure of the system, a broken sensor, incomplete readings, etc. Therefore, the inclusion of these data into the database requires special attention. Moreover, the manual inclusion of a great amount of data generated by the system is a very slow process. For instance, in a regular day of standard operation of the facility, 8000 curves can be measured, that is why it is virtually impossible to keep a daily control of the manual inclusion into the database.

The system has a feature for automatic inclusion of the measurement files. The design of this inclusion process is the result of a thorough experience and is being constantly updated with increasingly effective improvements, because this stage is remarkably important in order to avoid the inclusion of wrong data that can hinder the work of analysis.

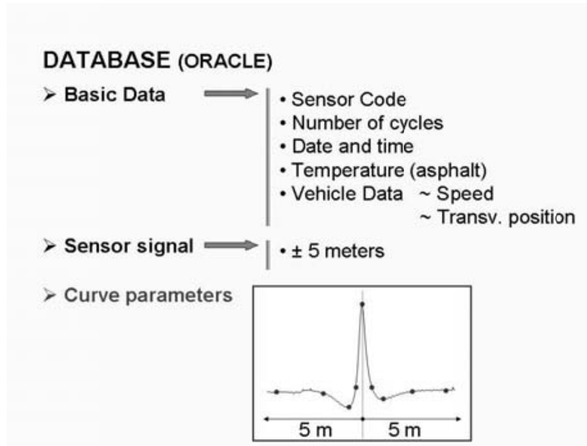


Figure 9. Registered data in a dynamic test

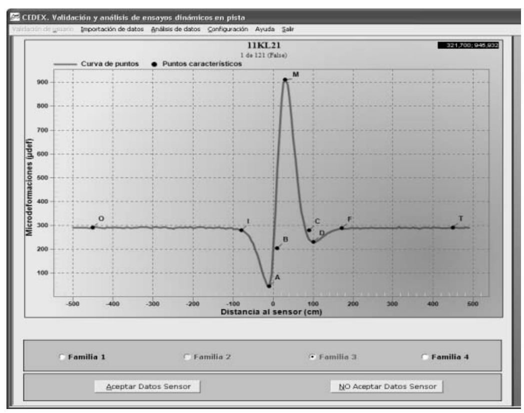


Figure 10. Example for the validation of a curve and its singular points

Prior to the data entry, the application displays the pavement temperatures assigned to each curve. The temperature at which the variable has been measured is very important in the case of pavements with asphalt mixture. If the temperature is incorrectly assigned, afterwards that will be difficult to identify, except when talking about wrong data, which will generate errors in the analysis stage. In the manual inclusion, there is no problem, because it is easy to detect it and it is corrected manually. In the automatic inclusion, the application is programmed to identify errors in the temperature measurements and it is able to correct them. This mathematical algorithm is based on a comparison by section and by stretch where the different sensors are placed.

Then, the program analyzes each type of curve measured and compares it with the type of variable that it is being measured. If there is a coincidence, the program calculates the singular points of the curve, as shows in Figure 13, and everything is stored in the database. In the automatic inclusion, if a difference arises between the measured curve and the expected typical curve, this curve is stored in a provisional file until the user evaluates it.

Curve Display

Every curve stored in the database can be displayed. Therefore, we can ensure the proper inclusion of these curves into the database; we can use it for instrumentation maintenance tasks and mainly to carry out the analysis of the measured variables.

To access the information, a form is required to be completed with additional data:

The type of sensor, the section, the range of transverse position, the range of vehicle speed, the range of temperature and the cycles period for which the records are required. The program then automatically displays the readings as can be seen in Figure 14, and the user can select any of the curves and print them as needed.

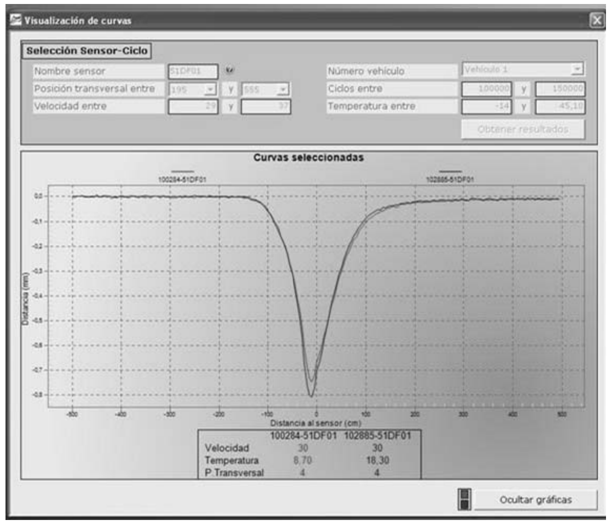


Figure 11. Example of display. Deflection at different temperatures

The restrictions when searching for curves are essential for a correct operation of the application, thus the application is opened to a large number of variations and that is why the user carrying out the analysis is also responsible for delimiting them depending on the required task.

Parameter Analysis

In the validation process of the curves measured by the sensors, the singular points are calculated for each curve. The database stores not only the curves and all the variables defining them, but also the coordinates of the singular points.

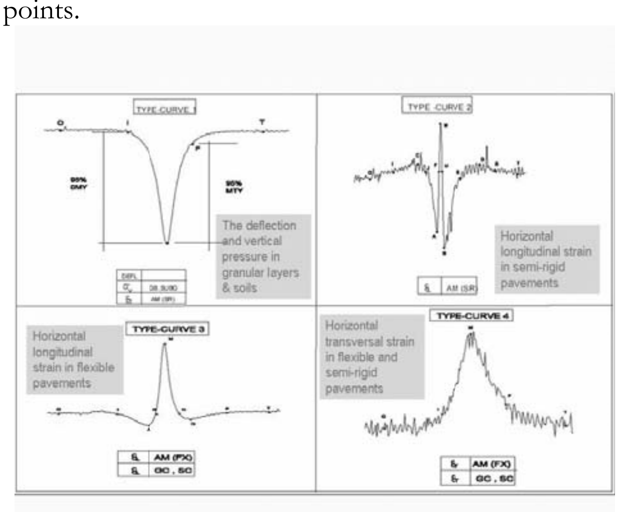


Figure 12. Singular points for the typical curves

In this part of the application, we can obtain graphical and numerical values for the different parameters versus the following variables: temperature, transverse position of the vehicles, vehicle speed and number of cycles. We define parameters such as the relative difference in the Y-axis or X-axis between two singular points of the same kind of variable. OMY value is typically used, since it represents the maximum peak response.

During the search process in the database, once the type of curve to be studied has been selected, the application displays the list of all the sensors of this precise kind. The next step is to select the parameter to be analyzed. Then, the system provides the ranges of temperature, speed, number of cycles and transverse position of the vehicle, in order to delimit the analysis as desired. The application provides the graphical evolution of the parameter with the selected variable. The search results can also be retrieved in an Excel file, in order to use the data in other computer applications.

The following figure shows a typical example of parameter variation with one of the variables:

- Type of sensor: Load cell.
- Variation relative to the pavement temperature.
- Parameter of search: In this case, OMY corresponds to the maximum vertical tension.
- Transverse position of the vehicle: 305, which corresponds to the central position in the transverse distribution of the vehicle movement.
- Between cycle 125000 and 200000.
- Vehicle speed: between 33 km/h and 37 km/h.
- Pavement temperature: between -3.9 °C and 40.5 °C. (The complete range with available data).
- Sensor: 52CT31, which corresponds to a load cell located at the top of the subgrade in section 5.

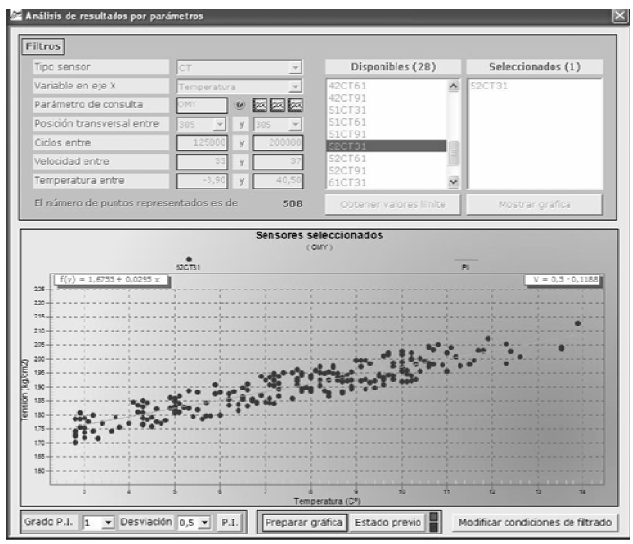


Figure 13. Example of parameter variation with temperature

Once all the values have been represented, the application offers the possibility to eliminate the outliers through a statistical procedure. For this purpose, the points are adjusted through a polynomial up to its third grade and we define a band consisting of that polynomial plus/minus the standard deviation multiplied by a number that ranges between 0.5 and 2. All the values that are out of this band will be eliminated.

Conclusions

- The facility is designed so that the traffic simulation vehicles move continuously, 24 hour a day, seven days a week. That is achieved by two control systems (micro computer and PLC) related to one another and both of them managed by a PC computer.
- The data acquisition system enables the full-automatic sensor measurement, based on different environmental and load related variables as well as number of cycles.
- The system allows the simultaneously managing of up to 256 sensors and typically, more than 250000 curves of the sensors are collected in a complete test.
- Records are stored in a Database together with the singular points of each one; e.g. maximum peak value, zero reference etc.
- A program enables the treatment of the stored data, so that the response of any sensor can be plot against test-related variables such as pavement temperature, vehicle speed or transverse position. The evolution versus the number of cycles can be displayed as well.
- Output ASCII files can be also generated so that data can be treated by a more specific software, like statistical packages or pavement design programs.

References

[1] "Sistema de adquisición de datos en tiempo real de la pista de ensayos acelerados de firmes del Centro de Estudios de Carreteras del CEDEX", B. Mingo, J. Pérez, J. Aleixandre, J. Bueno, J. Tamarit. CEDEX magazine "Ingeniería Civil" No. 114/1999. http://www.cedex.es/ingles/documentacion/publicaciones.html
[2] "Plan de instrumentación del 3º ensayo en la pista de ensayos a escala real del Centro de Estudios de Carreteras del CEDEX", J. Tamarit, J. Bueno, J. Aleixandre, B. Mingo, J. Pérez. CEDEX magazine "Ingeniería Civil" No. 115/1999. http://www.cedex.es/ingles/documentacion/publicaciones.html
[3] "Workshop: Introduction to Accelerated Pavement Testing", A. Mateos, N. Coetzee. "Impacts and Benefits from APT Programs" Accelerated Pavement Testing, III International Conference. http://www.cedex.es/apt2008

Stability Monitoring System Implementation for "Rural Mare Retezat" Dam

Susana Arad[1], Victor Arad [2], Joel Veres[3] ahd Timotei Neiconi[4]

[1]Assoc Professor, Electrical Engineering Dept., University of Petrosani, Universitatii str. 20, 332006 Petrosani, Romania, Phone: +40254542580, arad@dec.com.ro
[2] Prof., Mining Engineering Dept., University of Petrosani, 20 Universitatii str., 332006 Petrosani, varad@upet.ro
[3]Assoc Prof, Topography Dept., University of Petrosani, 20 Universitatii str., 332006 Petrosani, veresioel@yahoo.com
[4]Technical Manager, Hidroconstructia, Deva 2700-DEVA Aleea Florilor, Bl. 15, Tel.:054-214125; 214134;

Abstract

To avoid the hazards which can affect the construction stability, it is imposed to control the risk factors which can occur because of the geo-mechanical parameters modification in mass's rock and body dam's rock. The geo-mechanical parameter of rock samples was tested in laboratory and their variation will be analyzed and the crack of rock masses established. The paper takes into consideration the implementation of computer monitoring system and the optimization of network markers which are used in topographical measurements for tracing deformation purpose.

Keywords: Dam stability, monitoring system, remote control, rock mass properties

1. General Consideration

Gura Apelor dam is situated at the confluence of Lapusnicu Mare, Lapusnicu Mic and Ses rivers, where forming Raul Mare, at 47 km from Hateg City. The slope makes parts from mountain mass Retezat, Godeanu and Tarcu. Gura Apelor dam is situated at 600 m of confluence Lapusnicu with Ses River.

The upstream hydro-energetic development of Raul Mare River includes a dumped rock-fill dam which accumulates a water volume of 225 million m3, a pressure tunnel for the transport of water to the Retezat underground hydro-electric power plant with an installed power of 335 MW and a return power plant, the Clopotiva Power Plant. The production capacity of this power plant is of 629.5 GWh/year. The works began in 1975. The dam has a height of 168 m, and it is built of local materials: rock, filter and clay, Figure 1.

Figure 1. View on the Gura Apelor Retezat Dam

Evacuation of big waters is making through two gallery of emptying of the lake, and through discharging of big water, which is situated on the right versant. All the evacuations have a debit of 1700 m3/sec. Principal bringing workings provide transport of the water from the Gura Apelor to underground power station Retezat, consist from a tunnel with lengthwise of 18,432 m, excavated at a diameter between 5.5 and 6 m, with a final section of 4.9 m (interior concrete diameter).

The principal characteristics of Gura Apelor dam is: maxim high 168m above foundation, breadth at base 574m, total volume with filling 10,285,000 m3, rock fill 6,422,000 m3; clay in core 1,127,000 m3; filter and sorted gravel 183,000 m3; gravel 983,000 m3; filled in stability prism 883,000 m3.

The emplacement on the superior flow of the Raul Mare at just 660 m downstream of the Ses River confluence with Lapusnic River, which together forming Raul Mare, the rock-fill dam with a central nucleus of clay Gura Apelor create an accumulation with a total volume of 210 million m3 with 200 million m3 useful volume, representing the biggest accumulation from hydro energetic settlement of Raul Mare.

2. Geological and Geomechanical Characterization

General characteristics of Retezat Mountain are the presence of two big blocks of granite and granodiorite, one stronger developed on the principal north crest and other restrict on the principal south crest. Geological studies in the dam area, on the geologic elements base, structural-tectonics and mineralogical-petrography, have shown the difference between tree blocks (Bancila 1989).

It is known that Gura Apelor dam must be almost impermeable if is possible. The dam is known as non-homogeneous because there were used different materials (sand, gravel, clay) with different values of permeability coefficient, especially to reduce the infiltration debit, problem which represents a constraint for every hydro technique construction.

The Retezat Mountain has a lot of rivers, with a rich and permanent debit. The modification in time of precipitation and hydro energetic regime of versants, have a direct impact to shape the relief, and the tendency of rising the temperature determinate some modification of the vegetal cover.

From the hydro energetic point of view, on the natural open base, and in depth (drillings, gallery, caves), to realize a hydro geologic model of the dam, was determinate the characteristics: the level of underground water, permeability, water absorption, mineralizes of water and the consumption of solid materials necessary waterproof of rocks. From the geologic point of view, using geologic maps have tried to find a particular underground, but the underground need to be homogeneous, it don't need to present fracture or liquid accumulation which can produce the submerge.

Gura Apelor dam is situated in crystalline rocks and limestone belongs of Getic Clots, structures from granite, granodiorite, crystalline schist in the right versant and chlorite, gneiss in the left versant (Paunescu 2008). Rocks from the area are affected by two major faults and frequent discontinues. The characteristics of the realized section were taken from geologic documentation from Raul Mare. The lateral prism is composed from the same material with the upstream and downstream of the dam, with the characteristics shown in Table 1: Clay I (I), Clay II (II), Filter (III), Rock fill (IV) and Foundation (V).

Table1. Geotechnical characteristic of dam materials

Type of rock	I	II	III	IV	V
Unit Weight γ (MN/m^3)	0.019	0.02	0.02	0.025	0.0220
Friction angle φ (degree)	15	25	40	40	20
Cohesion c (MPa)	1	0.5	0	0	0

The dam is built from local materials, with clay nucleus. The rock from the emplacement foundation of the dam is built from hard granite schist, compacts and impermeable in the central zone and left versant. The dam prisms are made on the rock base with the exception in the river bed where is done on the existent gravel. The tight nucleus and the filters zone was foundation after removing the layer of rock of 3.5 m. In the nucleus zone was executed injection of consolidation on the rock, at 6-12 m depth. The left versant is especially made, on the superior part, of less resistant rock: peach stone, breccia, tectonical and altered. These rocks proved to be less permeable and injection works already done in the used technology give us more and more good results. The left versant zone - in the medium and superior zone, has been divided in two tectonic -structural blocks. As we observe in the studies we have made, that the formations of the two blocks are affected by cracks, which in time, at geological scale, have generated the alterations of the rocks changing them, in rocks which have lost the rocky characteristic.

The tight veil is built from two principal rows of horizontal drilling injected with 50-80 m depth and other two rows of drilling with 10 m depth in upstream and 16 m downstream.

Hoek – Brown failure criterion (Hoek 2002) is used to determine geo-mechanical parameters of rock mass. The criterion started from the properties of intact rock and then introduced factors to reduce these properties on the basis of the characteristics of joints in a rock mass. A rock mass damage criterion is introduced to account for the strength reduction due to stress relaxation and blast damage in slope stability and foundation problems.

Analysis of rock strength has been done by RocLab software developed to accompany this paper (Hoek 2002). This program provides a simple and intuitive implementation of the Hoek-Brown failure criterion, allowing users to easily obtain reliable estimates of rock mass properties, and to visualize the effects of changing rock mass parameters, on the failure envelopes. The results obtained from RocLab software are shown in Table 2. Joint properties and other structures have been estimated taking into account by direct shear test obtained in the Geomecanics Lab of Petrosani University (Arad&Todorescu 2006).

Table 2. Rock mass properties results from RocLab software

	Claystone	Granodiorite
Hoek Brown Classification		
σ_{ci} [MPa]	35	175
GSI	38	43
m_i	4	29
d	0.7	0.7
Hoek Brown Criterion		
m_b	0.132616	1.26545
s	0.000125211	0.000258434
a	0.51302	0.509269
Failure Envelope Range	Application_Slopes	
σ_{3max} [MPa]	2.26079	3.55459
Unit Weight [MN/m3]	0.02	0.025
Slope Height [m]	168	168
Mohr-Coulomb Fit		
c[MPa]	0.284633	1.3642
Φ [degree]	20.5604	48.9267
Rock Mass Parameters		
σ_t[MPa]	-0.0330458	-0.0357391
σ_c[MPa]	0.348402	2.6059
σ_{cm}[MPa]	1.58324	25.3128
Em[MPa]	455.124	714.151

The results obtained by lab tests data analysis and by simulation with RocLab software allow us to predict the behavior of the dam body and hydro energetic construction state and to realize matrix with industrial risk.

3. Slope stability analyses

3.1 State of Art

To monitoring of the dam behavior was take in consideration 33 markers embedded on the dam, disposed on four alignments in downstream and three alignment in upstream, Figure 2. We realized the optimization of network markers which are used in topographical measurements for tracing deformation purpose (Dima 2004).

Equipment for the behavior control of construction it achieved and variety in the same time with the execution of investment. These are: geologic and piezometer-type drillings; measuring and control devices (MCD); pendulum; topogeodesy equipment for determination of movement in horizontal plane, vertical and in space; seismic station, transducers and automatic tracking systems for controls of signals (stress, strain, temperatures and electrostatic field).

This works was done on tree specific area of the left slope, named: (i) Zone I-Schist, (ii) Zone II – Breccia and (iii) Zone III-Granite.

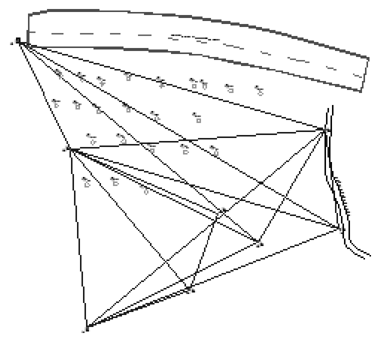


Figure 2. Monitoring markers outline

The measurements data, the experimental results from laboratory and all the data for monitoring activity are realized by specialized team or experts in domain. Drilling and injection works at Gura Apelor Dam was started in the year 2004 with execution of experimental drilling realized with classic endowment and have continued with a new lot of drilling executed with modern technology, Figure 3.

Supervision of the rocks deformation in foundation is done with the help some telemetry devices. The rotation or inclination of surface from dam body is measured with inclinometer. The extensometer is used for specific deformation. At all the instruments for resistive measure, the measurement is done from afar with the help of the telemetry (Arad&Arad 2006).

To have safety data for the dam measurement are necessary to obtain the external and internal stress, horizontal and vertical displacement over the fissure, interstitial pressure and so on. The inner temperatures are measured with the help of resistive or electric –acoustic remote control devices, the measurement device are situated in the section or point, the measurement is done at distance (downstream parameter, visitation gallery) through some electric cables. Specific deformation is measured with electric, acoustic or resistive extensometer.

Slope stability is put in danger by destroying the local equilibrium between forces which stress the slope and interior strength forces of rocks, under the direct action of diverse internal or external factors, naturals or artificially. Loosing of slope stability is producing through deformation and sliding of slope, on a surface, because of loosing equilibrium limit of rocks, expression through minimum stability coefficient.

Figure 3. Aspects from the waterproof process

3.2. Analysis of the slope stability

Slope stability is endangering through disturbing the local or ensemble equilibrium between the forces which solicit the slope and internal resistive forces, under the direct action of diverse internal or external factors, naturals or artificially. Loosing slope stability is producing through deformation and sliding of those, after a surface, because exceeding equilibrium limit state of rocks shown through minimum stability coefficient.

To eliminate slides of the field and trickling of water from left slope was performed a lot of drilling. From this drillings was established loosing of water from versant, and after was realized a veil of tight thru injection of water-cement- bentonit in massif.

TALREN software is ideal for checking the stability of geotechnical structures, with or without reinforcements: natural slopes cut or fill slopes, earth dams or dikes. In the present version of TALREN, the safety factor Γ is calculated by TALREN which should be ≥ 1 for equilibrium. TALREN used method of the limit equilibrium calculation along potential failure surfaces using the Fellenius method. The value of Γs, Γc and $\Gamma\varphi$ is imposed for each soil of the geometric model like in Figure 4, based on geotechnical characteristic of dam materials, shown in Table 1.

For the dam model with a full lake, with prism, friction internal angle for foundation $\varphi = 30$, was obtained the minimum safety coefficient $\Gamma = 1.55$ for the sliding surface from the Figure 4.

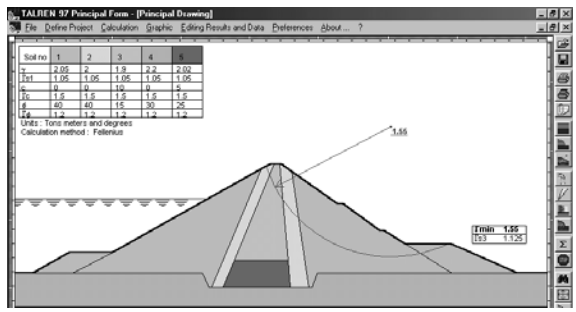


Figure 4. Safety coefficient determination

There is often in rock engineering to establish the geology and structure of the rock mass remotely. In this situation the choices are traditional photogrametry techniques enhanced by the advances of digital photography and an adaptation of the advances of digital airborne laser technology. Analysis of the rock slope requires a geometric description of the surface and knowledge of fracture patterns and their properties.

The geographic information system (GIS) is increasingly viewed as a key tool for managing spatial distribution of data (Grecea 1999). Advanced in digital technology provide tools that allow the generation of digital terrain models at the centimeters scale. The photogrammetry survey used two camera stations located on the upstream and downstream edge of the dam forming the dam crest.

Natural and human-induced slope movement and slope failures are complex geotechnical engineering problems involving both surface and subsurface conditions and their interactions to triggering factors. Current geotechnical modeling tools are focused on the numerical analyses and generally not designed to facilitate the requirements of site investigation and characterisation. The geophysical and geotechnical model of the dam and of the slope will be made with the data measured from monitoring the dam and the determined parameters, this model being integrated in the monitoring system.

4. The Monitoring System of Dam Stability

The architecture of the integrated system is presented in Figure 5 (Calarasu 2008). The system is composed of three stages which have to be followed for the implementation of specific tasks, like in Table 3.

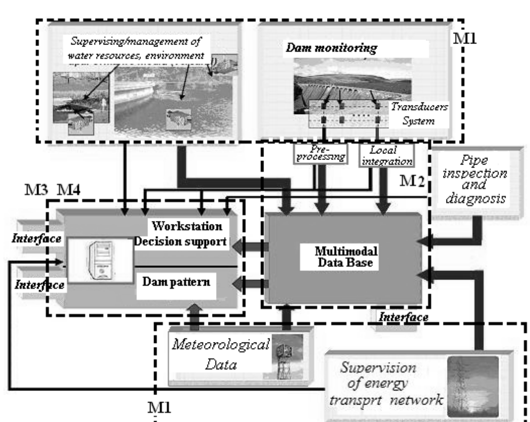


Figure 5. Architecture of the integrated system

Table 3. The stages on the implementation of integrated system

Stage 1: Data collection, Management & Synthesis		
Input data	Synthesis tools: GIS	Output
➢ Base Digital Elevation Model Data ➢ Geology/Geophysics ➢ Displacements ➢ Groundwater ➢ Field Observations	Field ↓ Office ↓ Central Spatial Data Base	Data integration Data presentation Data vizualization Data comunication Ground modeling
↓		
Stage 2: Development of Engineering/ Geotechnical model		
Input data	Tools: CAD software	Output
➢ Ground model Stage 1 ➢ Material properties ➢ Pore pressure		Geotechnical model
↓		
Stage 3: Engineering/ Stability analyses		
Input data	Tools: Numerical analysis	Output
➢ Geotechnical model ➢ Material properties ➢ Pore pressure		Engineering decisions

The integrate system contains many moduli, grouped in four functional categories (Figure 5), which can be explained as in Figure 6:

- Monitoring dam parameter, data acquisition and transmission (M1);
- Conceptual organization and pre-processing the data (M2);
- Post- processing data (M3);
- Data interpretation, assisting decisions (M4).

The monitoring, getting and transmission of data is made by means of automation equipments and the related software (firmware).

The group of conceptual organisation of data and preprocessing contains modules/software components which undertake the logging from transducers and operate and put them in the data base.

The modules from the preprocessing zone deal with data extraction according to the configurations of the hydrotechnic facilities and the machine work requirments.

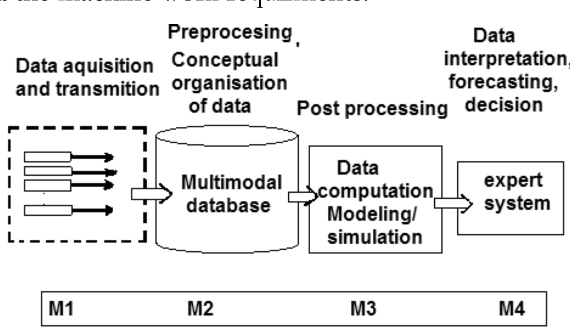


Figure 6. Block diagram of system

The functional blocks from the execution zone /decision assisting implement the computerised visual inspection, modelling/simulation/prediction functions, expert systems.

From identification the risk factors was established the category of elements necessary in dam monitoring for construction safety and to avoid the risk apparition:

- Morphological and hydro geological changes on the lake versants zone of dam emplacement.
- Changes in strength and tight structure of dam body.
- Hydro mechanic equipment function (empty bottom).

The systems functions of measuring and control propose have in consideration (measuring and control device): the water level in the lake; size and duration of rainfall; groundwater level downstream dam; interstice pressure; total pressure; linear specific stains; deformation rocks left side; fillings settlement; face and crowning deformation; closing - the opening of joints in left side mining working; geophysical characteristics of the foundation rocks; seismic control; warping and erosion; direct observations.

The parameters measured, which represents the input of the system, are: relative displacements measurement of the dam, the variations in the level of the water surface, ex-filtration, temperature and distance measurement, stress and strain in body and side of dam.

The system proposed allow the management of the specifically the data base about visual inspection of hydro electric settlement, with a specialized component for behavior analysis of the wall dam based on, computer vision' technical.

It is necessary to find a probabilistically model with data base measurements joint with the laboratory tests, on statistics criteria which should estimate the prediction of conduct in order to determine the normal or abnormal state of behaviour in the analyzed construction.

5. Conclusions

Based on investigations and a monitoring program, Hidroconstructia Company (RMR Hydro) decided that it is neccesary to be controlled the structuraly stability of Gura Apelor dam and to improve it. It was given special attention because of the serious consequences which could have resulted from the danger of slope slides or the dam body failure.

The displacement in different points of construction, measured with the measuring devices, show the answer of the building at external or internal stress.

The photogrammetry survey used two camera stations located on the upstream and downstream edge of the dam forming the dam crest.

Current geotechnical modelling tools are focused on the numerical analyses and generally not designed to facilitate the requirements of site investigation and characterisation.

The system proposed allow the management of the specifically the data base about visual inspection of hydro electric settlement, with a specialized component for behaviour analysis of the wall dam based on ‚computer vision' technical.

References

[1] Arad, S., Arad, V., Lupu, L., (2006). „The remote control of Electric Devices, Protection, Control and Monitoring Systems", Proc. of Int. Symp. ISIE 2006, IEEE Conference Proceedings, vol I, pp. 267-272, 0018-9219

[2] Arad, V., Todorescu, A., (2006). „ Ingineria rocilor si structurilor de suprafata", Risoprint Publisher, Cluj Napoca

[3] Băncilă, I. (1989), Geologia amenajărilor hidrotehnice,Tehnical Publishing House, Bucureşti

[4] Calarasu, A., Stoian, I., s.a (2008). Integrated system, decisional support based on multisensorial information fusion for behavior surveillance and prediction of dams and hydro-power plants. Proc. of Int. Symp. AQTR 2008, Cluj Napoca

[5] Dima N., Herbei O., Veres I., (2004). Optimisation of geodesical and topographycal networks Annals of the University of Petrosani 2004 1454-9174 pp.14-21

[6] Grecea, C., (1999). „Introduction in satellite geodesy". Mirton Publisher, Timişoara

[7] Hoek, E., Carranza-Torres, C., Corkum, B. (2002) Hoek-Brown failure criterion – 2002 edition. Proc. of North American Rock Mechanics Soc., Toronto.

[8] I.S.P.H. Bucureşti, (1979). Calculul privind stabilitatea barajului Râul Mare Retezat, Hidroconstructia, Hateg

[9] Paunescu, N. (2008).„Geologic and geotechnic study about rocks from Raul Mare area", Petrosani University

[10] TALREN software <http://www.terrasol.com>

An Advanced Wireless System for Emergency Management in Hospitals

Angelo Biscotti[1], Sara Spadoni[1], Berardo Naticchia[1] and Alessandro Carbonari[2]

[1]Polytechnic University of Marche, Smart Space Solutions SRL, Via Brecce Bianche, 60131 Ancona, Italy
[2]Polytechnic University of Marche, DACS Department, Engineering Faculty, Ancona, Italy. Tel: +390712204397; Fax: +39 0712204582; Mobile: +39 3402502768, alessandro.carbonari@univpm.it

Abstract

This paper reports preliminary results achieved during the experimentation of a new real time localization system, especially designed in support of hospitals' emergency management. This system represents the first step of a wider ongoing research project, carried on by the authors in collaboration with the academic spin-off company Smart Space Solutions SRL, aimed at the developing of the W.E.M.A.S. (Wireless Emergency Management Aiding System) system, whose final aim is increasing hospitals' efficiency in the management of the maxi emergencies consequent to disasters. WEMAS uses advanced wireless localization tracking and communication technologies to provide a number of features, such as: patient's flows and patient's healthcare progress tracking, equipment and medical personnel tracking, triage information management, ambulance scheduling and early communication of patient status from ambulance to the receiving emergency division. As a side effect, WEMAS system will equip hospitals with this new network infrastructure to optimize patient logistics for routine management. This paper describes also tests performed on the first release of WEMAS system, in the Emergency Division of the Hospital of Senigallia (Italy), demonstrating its features and technological feasibility.

Keywords: RTLS; wireless communication; disaster response; emergency management.

1. Introduction

This paper concerns the first step of a wider ongoing research project, named WEMAS (Wireless Emergency Management Aiding System), and carried out by the Department of Architecture Constructions and Structures (DACS) of the Polytechnic University of Marche, the academic spin-off company Smart Space Solutions SRL and the hospital of Senigallia.

The WEMAS, that will be developed to increase hospitals' emergency divisions efficiency in the management of maxi-emergencies consequent to disasters, in its final release will provide the following functionalities: support to first aid management on disaster fields; support to patient transfer from disaster field to hospital; support to emergency division activities within hospital.

As a consequence, the WEMAS will have positive implications also in hospitals' daily activities, allowing patients' healthcare flow tracking, optimization of resource management and also facility tracking to prevent loss or robbery attempts.

The project is compliant to the Marche Region's legislation that, by means of the regional act "DGR 49/2004", imposes that every hospital must arrange a PEIMAF plan ("Piano di Emergenza Massiccio Afflusso Feriti") devoted to coordination of hospital's activities in case of massive victim afflux consequent to disasters.

Nevertheless, empirical experience has attested the inadequacy of the PEIMAF plan in maxi-emergencies management, because of the adoption of traditional communication and management instruments. Even the most recent research, carried out in countries very sensitive to catastrophic events, like Japan or USA, realized inadequacy of traditional means when coping with catastrophic events (Wickramasinghe et al. 2006). The last catastrophic event is represented by September the 11th attacks, when inefficient communication between disaster field and neighbor hospitals was experienced.

Nowadays, limits could be fortunately overcame by the advent of recent IC technologies, providing a valid support in hospital's activities management (Chao et al. 2007), like faster patient's identification, patients' healthcare progress real-time tracking and so on. Some examples of such systems have been developed, tested and adopted in hospitals, but they all tend to be very invasive, moreover when installed in existing buildings.

This paper presents first studies, aimed at developing a first WEMAS system release for emergency divisions, carried out by the authors through the following steps:

- performance analyses to discern the system's main functionalities;
- first localization experimental tests, to perform the reliability of the system, through the use of signal quality indices (these tests have been carried out in one of the most difficult environments: the hospital of Senigallia).

The system will be based on ZigBee technology, as it provides such a tiny structure, considered critical for fast and reliable communication.

2. State of the Art

PEIMAF standard plans organize rescues operations on a three level structure: organization of a field hospital for first aid; victims transfer form disaster scene to hospital; victim management and care within hospital's divisions.

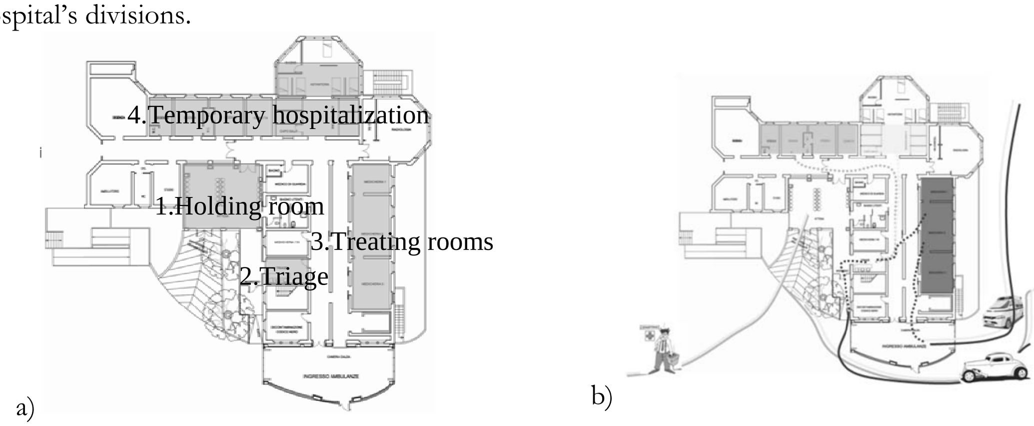


Figure 1. Senigallia Hospital Emergency Division organization in daily routine (a) and re-organization of victims paths in case of disaster management (b).

PEIMAF is applied just in case of massive patient flux and requires to fully re-organize rescue operations, from their triage to further medical care (Figure 1). As it can be noticed from comparing the two parts of Figure 1, in case of disaster Senigallia hospital emergency division occupies new areas, involving them in victims' care process. Figure 1-b also shows that victim's flows follow different courses in relation to their triage code. Due to the involvement of all hospital divisions, it requires a constant and continuous communication between personnel, not achievable through standard communication protocols.

The introduction of Radio Frequency Identification technologies could give a great impulse to the development of innovative real-time management systems, applied at different complex operative environments like hospitals. These systems, based upon various communication protocols (WiFi, IRDA, RFID, UWB) have been installed and employed in few hospitals. However they all present the same limits, such as:

1) high power consumption, requiring connection to the electric system;
2) invasive installations of big sized devices and cabling.

For that reason, the WEMAS project wants to release a new product.

3. The New Approach Brought about by Our W.E.M.A.S.

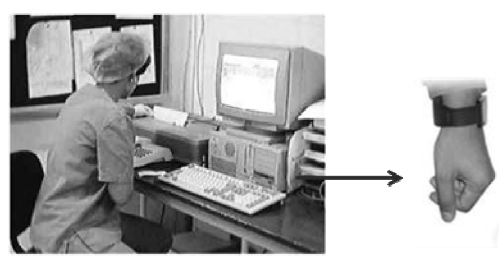
Table 1 illustrates the results of analyses carried out together with the hospital of Senigallia's personnel, relative to all the functionalities that could be reasonably provided by the WEMAS for automating hospital practices. For instance, check-in would change it standard procedures (assignment of a paper form to every patient containing his/her ID, triage code and health status), in favor of a badge to be automatically updated during patient's healthcare progress. The main limit of the standard practice lays in the frailty of paper documents and risks of document loss.

Instead, according to the WEMAS, patient's ID and data will be recorded in the central hospital database through a local computer unit or a PDA, passing through emergency division's intra-network. The same ID will be also associated to a bracelet worn by the patient and containing a Zigbee End-Device dispositive, able to communicate with portable and fixed wireless devices of a network installed in the hospital. Portable devices will also be provided to doctors, nurses and hospital assets, allowing for real time monitoring of patients and healthcare resources' positions.

Table 1. The WEMAS' features foreseen for the final release.

Query	***Check out***
- *Patient's identification* - *Healthcare plan* - *Clinical parameters* - *Real-time location* - *Timeling* - *Available resource*	- *Admission* - *Dimission* - *Self-dimission* ***Health care*** - *Triage code assignation*
Alert	***Request***
- *Patient's support call* - *Patient's clinical parameter monitoring* - *Patient's drift* - *Robbery attempts* - *Healthcare progress tracking*	- *Medical expertise scouting and consulting* - *Medical test or instrument reservation* - *Busy* - *Call monitoring*

In addition, by means of their PDAs, doctors will be able to visualize patient's identity and his healthcare progress, by connecting to patients' bracelets. Through their devices they could also ask for a medical expertise scouting and consulting, as shown in Figure 3.


a)


b)

Figure 2. Traditional triage paper (a); Programmation and Assignation of patient's identification bracelet (b).

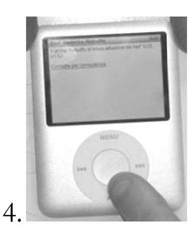
1. 
2. 
3. 
4. 
5. 

Figure 3. The procedure for "Medical expertise scouting and consulting" through medical PDAs

4. The WEMAS Technology

The WEMAS provides its features through hardware and software components based on the IEEE 802.15.4 standard medium access and Zigbee stack communication protocol. These protocols are intended for use in embedded applications requiring low data rates and low power consumption, and present a lot of useful properties to be implemented in health care workplaces (Lorincz et al. 2006).

ZigBee's current focus is to define a general-purpose, cheap, self-organizing mesh network. There are several advantages provided by the Zigbee protocol, that make it strategic for hospital environments: it works in ISM band; it is reliable and self healing; it supports large number of nodes (16 bit address space); it is easy to deploy; it allows very long battery life; it is a safe and low cost technology.

The WEMAS exploits three Zigbee device types (Figure 4-a), compliant to standard application profile: each one presenting specific features, as depicted in Table 2.

Our WEMAS devices are equipped with the integrated radio and microcontroller Chipcon/Texas Instruments CC2431 chip with 128K flash memory. CC2431 has an embedded localization hardware engine, which estimates the position of a node by evaluating the received signal strength (RSSI) attenuation (Michel et al. 2006), according to the Path Loss Model (PLM) formula:

$$RSSI = -\left(A + 10 \cdot n \cdot Log_{10}(d)\right) \quad \text{(eq. 1)}$$

where d represents the distance in meters from the sender node, while A is the received signal in dBm at a distance of one meter with a sending power of 0dbm.

The decentralized localization algorithm is performed entirely by the “blind” node, reducing the communication costs and bottlenecks. A mobile node who wants to evaluate its position sends a broadcast message to all ZRs of the network in its radio range; each ZR replies with a short message whose payload contains its fixed coordinates and the message’s RSSI received, measured in dBm (see Figure 4-b). At least three RSSI indications are then written in the location hardware registers with ZRs coordinates, A and n values of the PLM formula in eq. 1. Output contains the coordinates of the evaluated position. The procedure to evaluate A and n values is described in (Chipcon/Texas Instruments 2007).

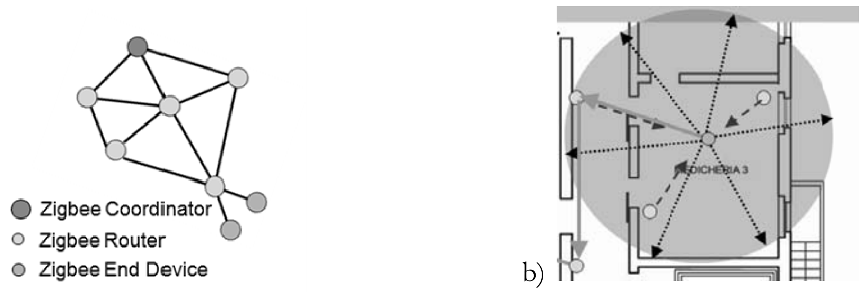
Table 2. ZigBee device types with their features

Zigbee Coordinator (ZC)	**Zigbee Router (ZR)**	**Zigbee End-Device (ZED)**
• one and only one required for each ZB network; • initiates network formation; • acts as 802.15.4 PAN coordinator (FFD); • not necessarily dedicated device, can perform applications.	• may associate with ZC or with previously associated ZR; • acts as 802.15.4 coordinator (FFD); • local address (destination) allocation/de-allocation; • participates in multi-hop routing of messages.	• mobile node; • shall not allow association; • shall not participate in routing; • put to sleep by parents.

The engine is not capable to evaluate node’s position if less than three ZRs reply to the broadcast request. In this case the position can be evaluate through gate control. The utilization of a dedicated hardware reduces considerably the delay of the mathematical process. The Zigbee standard profile suggests to have main powered ZR, in order to maintain the radio transceiver in listen mode and have the possibility of routing messages throughout the network. The necessity of using cables dramatically decreases the flexibility of the network in existent hospital environments, because fixed nodes have to be placed in proximity of an electrical source. However the Zigbee specifications allow to use a synchronous communication protocol which uses a special type of message called *beacon*.

The beacon is a sync message, which defines a time interval for message passing, but increases extremely the latency, in order to minimize duty cycle. WEMAS system uses a new approach, developed by the academic spin-off Smart Space Solutions srl. This proprietary protocol is intended for reduce ZR’s duty cycle in asynchronous mode. It allows the arrangement of full wireless networks, because battery powered, requiring minimum maintenance efforts, with a 2 years long power life.

Each ZR node remains in sleep mode for reducing power consumption until it has to route a message or reply to a request. In these cases an asynchronous signal wakes it up from sleep and allows it to work without waiting for a sync.



a) b)

Figure 4. example of a mesh Zigbee network (a); RSSI based message exchange for localization (b).

The localization hardware engine and the low power communication capabilities, combined with the self-healing property, make the WEMAS extremely suitable for hospitals with respect to other tracking technologies.

RFID has a lower range, compared with costs of readers, and interference problems with existing LAN networks. Zigbee compliant devices instead aren't subject to significant interference with other wireless technologies. In (ZigBee Alliance 2007) it is showed how in addition to DSSS (direct sequence spread spectrum) technology, the 802.15.4 physical layer protocol increases the opportunities for coexistence by employing a technique, generally known as frequency division multiple access (FDMA). Even with the techniques described above, a ZigBee device may find itself sharing a channel with interferers. The approach taken by the IEEE in the 802.15.4 standard is known as carrier sense multiple access (CSMA). Both tests and everyday use in real environments with real data traffic bear prove ZigBee's robustness.

Infrared systems need to have a *line of sight* to communicate one another and have not complex communication capabilities, so it can be used just for gate control. Instead Zigbee devices do not suffers multipath effects, thanks to application of mesh networking. Since multiple paths exist in a mesh network, messages can be sent into several directions. If one path is lost for some reasons or the link budget of a connection is too low, the message is likely to direct it along another path.

UWB allows very accurate positioning, but it is very expensive and complex.

Wi-Fi boasts a higher range, but the infrastructure is expensive and needs electric power. The signal strength of a Zigbee device is proportional to the transmission power and decreases in presence of obstructions such as doors or walls (Figure 5).

Some environments need to increase the number of ZR to maintain redundancy and avoid bottlenecks.

Factor	433 MHz		868 MHz		2.4 GHz	
	Loss	Attenuation	Loss	Attenuation	Loss	Attenuation
Open office	0 %	0 dB	0 %	0 dB	0 %	0 dB
Window	< 5 %	< 1 dB	15 %	1 – 2 dB	30 %	3 dB
Thin wall (plaster)	25 %	3 dB	35 %	3 – 4 dB	50 %	5 – 8 dB
Medium wall (wood)	40 %	4 – 6 dB	50 %	5 – 8 dB	70 %	10 – 12 dB
Thick wall (concrete)	50 %	5 – 8 dB	60 %	9 – 11 dB	85 %	15 – 20 dB
Armoured wall (reinforced concrete)	70 %	10 – 12 dB	80 %	12 – 15 dB	90 %	20 – 25 dB
Floor or ceiling	50 %	5 – 8 dB	60 %	9 – 11 dB	85 %	15 – 20 dB
Armoured floor or ceiling	70 %	10 – 12 dB	80 %	12 – 15 dB	90 %	20 – 25 dB
Rain and/or Fog	90 %	20 – 25 dB	95 %	25 – 30 dB	?? *	?? *

* = *Attenuations increase along with the frequency. In some cases, it is therefore difficult to determine loss and attenuation value.*

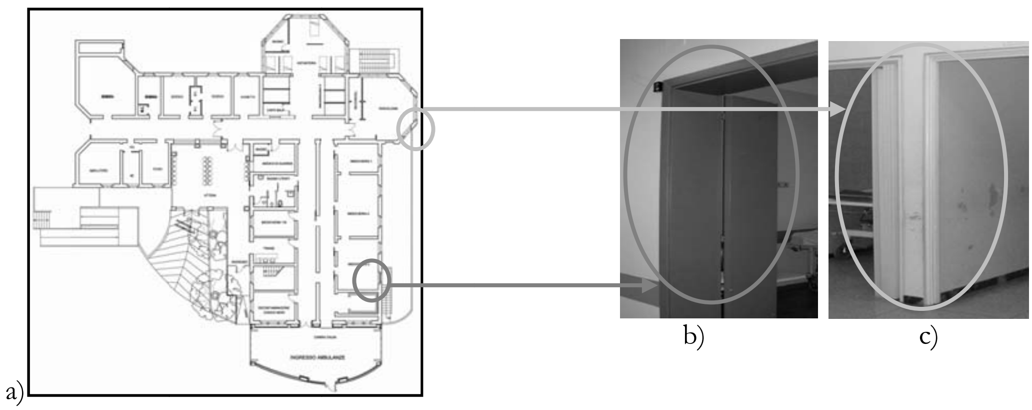
Note = *The table above is only indicative. The real values will depend on the installation environment itself.*

Figure 5. Signal attenuation in correspondence of different kinds of obstructions (OneRF Technology, 2007).

5. Preliminary Experiments on the Wemas Real-time Localization System

As previously explained, in order to assure that all the WEMAS' features will work properly, it is necessary to develop a reliable wireless sensor network (i.e. reliable communication among all the nodes that is measured through RSSI received by every ZR within radio range).

Preliminary tests were performed in the worst environment where WEMAS could be likely installed, that is to say the emergency division of the Senigallia hospital, built with a heavy masonry structure (see Figure 6-a), whose thickness ranges between 0.4 and 1.2 *m*.



a) b) c)

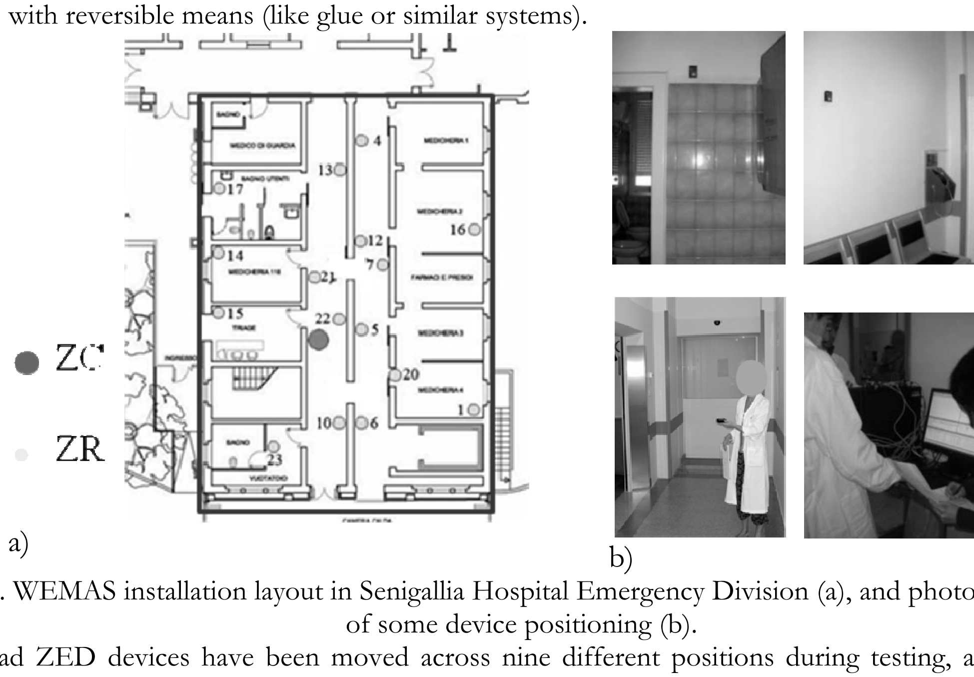
Figure 6. Ground level plan of the Senigallia Hospital Emergency division (a); details of a 0.4 m thich solid brick wall (b) and of a partitioning wall (c).

Senigallia Hospital Emergency Division is a three level building, characterized by very thick solid brick made walls (Figure 6-b), and partitioning walls built with hollow-solid bricks (Figure 6-c). Its cramped space organization makes localization a tougher task.

Following first preliminary experiments, the final testing have been carried out at the ground level of the Emergency Division. Figure 7 illustrates fixed devices positioning throughout the structure: as it can be noticed a ZC device has been placed close to the middle of the left side corridor, next to the left wall, while sixteen ZR devices have been deployed as follows:

- nine ZR in the two corridors (no. 4, 5, 6, 7, 10, 12, 13, 21, 22) ;
- six ZR in six different rooms along the corridors (no. 1, 14, 15, 16 17, 23);
- one ZR (no. 20) next to the door of one room.

The photographic report in Figure 7-b depicts some of the routers installed on the hospital's walls: it can be noticed that this suggested technology requires rather small sized devices, properly packaged and fixed on the walls with reversible means (like glue or similar systems).



a) b)

Figure 7. WEMAS installation layout in Senigallia Hospital Emergency Division (a), and photographic report of some device positioning (b).

Instead ZED devices have been moved across nine different positions during testing, as illustrated in Figure 8-a, in order to verify that all the room included within the emergency division can be tracked by our system. Localization have been performed using the algorithm detailed in paragraph 4.

It has to be outlined that our sensor devices hold a built-in mechanism for RSSI evaluation of the received signal always limited between -95 dBm and -40 dBm (that is to say, if the received RSSI is lower than -95 dBm than it is neglected; otherwise if it is higher than -40 dBm, than it is outputted at -40 dBm). A good communication quality is assured if RSSI value higher than -85 dBm is measured by at least one ZR, when the ZED is located in every position of the building.

Experimental tests gave back good results from this point of view. Figures 8-b and 8-c show some examples of the performed localization tests: in every case the received RSSI was higher than -85dBm, the strongest ZR was identified as the closest and a good match between the real and the estimated position for the blind node (i.e. ZED) was obtained.

As guide-line for further installations it was also noticed that: a 350 m² large built area, surrounded and subdivided by heavy walls, can be monitored by using 16 ZRs, that is to say one ZR per 20 m² large built area. ZRs must also be installed next to all door openings, in order to avoid ZED localization errors.

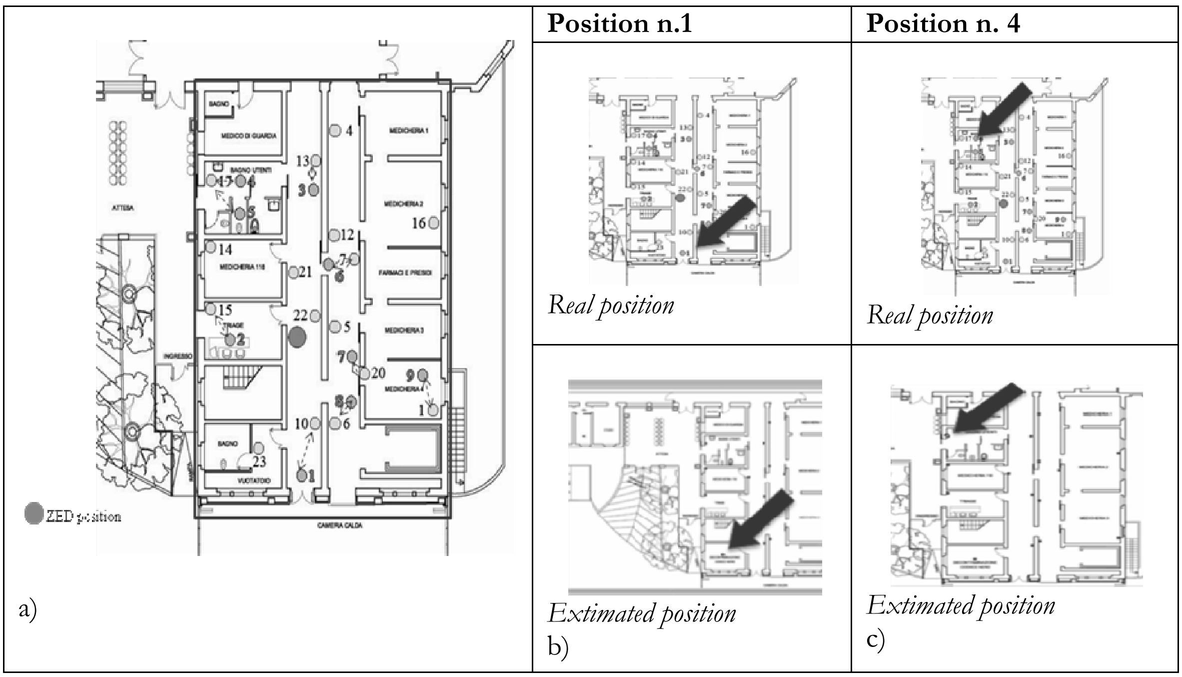


Figure 8. Various locations occupied by the ZED devices during testing (a); comparison between real and estimated positions (b, c).

6. Conclusions

Performance analyses relative to maxi-emergencies management allowed identifying a series of features that our WEMAS can offer in order to become a strategic instrument either in the case of massif victims affluence, or in hospital's daily activities. In the last case, the adoption of the WEMAS can be useful for optimized resource and communication management, besides improving structure's safety level, achievable through the installation of the full-wireless, low invasive system, based on our unique technology.

First localization experiments demonstrated the reliability of the WEMAS, even in case of buildings having very heavy structures.

Within little time, a prototypal version of WEMAS system, endowed with some further functionalities of Table 1, will be installed and tested in the emergency Department of Senigallia Hospital for continuous experimental monitoring.

References

[1] C. Chao, W. Chen, Y. Chy, B. Lin (20079, "Improving patient safety with RFID and MOBILE technology", Int J Electron Healthc, 3(2): 175-92

[2] Chipcon / Texas Instruments (2007), "System-on-Chip for 2.4 GHz ZigBee®/ IEEE 802.15.4 with Location Engine" <URL:http://focus.ti.com/lit/ds/symlink/cc2431.pdf>

[3] K. Lorincz, M. Welsh (2006), "MoteTrack: a robust, decentralized approach to RF-based location tracking", Pres Ubiquit comput., Springer-Verlag London Limited, DOI 10.1007/s00779-006-0095-2.

[4] D. Michel, D. Toggenburger (2006), "Sensor network soccer", Diploma Thesis WS, University of Applied Sciences Rapperswil HSR, Department of Electrical Engineering, Institute for Communication Systems.
[5] OneRF Technology (2007), “B2400MC-uTiny Technical Manual” <URL:http://backoffice.inware.it/files/prodotti/download/onerf/manuale%20zigbee%201.pdf>
[6] N. Wickramasinghe, R.K. Bali, R.G. Naquib (2006), "Application of knowledge management and intelligent continuum for medical emergencies and disaster scenarios", Conf. Proc. IEEE “Eng Med Biol Soc.”, 1:5149-52
[7] Zigbee Alliance White Paper (2007), "ZigBee and Wireless Frequency Coexistence", < www.zigbee.org/imwp/download.asp? ContentID=11745>

Development of an Automated Safety Assessment Framework for Construction Activities

Seokho Chi[1] and Carlos H. Caldas[2]

[1]Ph.D. Candidate; Dept. of Civil, Architectural and Environmental Engineering, University of Texas at Austin, TX, 78712; (512) 471-7950; shchi@mail.utexas.edu
[2]Assistant Professor; Dept. of Civil, Architectural and Environmental Engineering, University of Texas at Austin, TX, 78712; (512) 471-6014; caldas@mail.utexas.edu

Abstract

This paper presents an ongoing research project concerning the development of an automated safety assessment framework for earthmoving and surface mining activities. This research seeks to determine data needs for safety assessment and investigates how to utilize collected data to promote more informed and efficient safety decision-making. The research first examined accidents and fatalities involved with earthmoving and surface mining activities—more specifically, those involving loading, hauling, and dumping operations,—investigated risk factors involved with the accidents, and finally identified data needs for safety assessment based on safety regulations and practices. An automated safety assessment method was then developed using the data needs that had been identified. This research is expected to contribute to the introduction of a fundamental framework for automated safety assessment and the systematic collection of safety-related data from construction activities. Implementation of the entire safety assessment process on actual construction sites remains a task for future research.

Background and Motivation

A construction site is typically equipment-intensive, and more than a thousand workers die every year in the United States due to heavy-equipment-related fatalities. The Bureau of Labor Statistics (BLS) analyzed primary and secondary sources of machinery-related fatalities from 2003 to 2006 based on heavy equipment types (BLS 2003-2006). They reported that 1,644 (34%) of the total 4,796 fatalities resulted from operation of the six common types of heavy equipment machinery: excavating machinery, loaders, road grading and surfacing machinery, cranes, trucks, and forklifts. Given these fatality numbers, many researchers have looked at positive ways to achieve safer working environments, trying to identify risks and safety hazards on a site. Questions about how safe a site is or what kind of safety culture already exists depend upon how many risk factors exist in construction activities, and site safety assessment is a pre-requisite to identifying such risk factors that contribute to accident potential that need to be controlled (Ahmad and Gibb 2004). unacceptability of safety conditions (Ahmad and Gibb 2004). In addition, such human observations are time-consuming, and it is almost impossible for observers to monitor site safety at all time; accidents are likely to arise suddenly.

For these reasons, much research has been performed for automating the safety assessment process. Although these studies have made an effort toward safety improvement, most of them have merely presented site information acquisition and processing techniques without providing in-depth explanations as to what kinds of data are required for safety assessment or how data is related to construction accidents. For example, some studies asserted that the on-site tracking of objects is important for safety assessment, but they did not show which types of accidents were prevented by object tracking and how the object tracking was used for safety assessment. Thus, we have a research need for the development of a step-by-step safety assessment framework covering activity risk and data need analyses to safety decision-making.

The primary purpose of this research is to develop an automated safety assessment framework for construction activities. The course of this research began with a literature review on earthmoving and surface mining activities, specifically, loading, hauling, and dumping operations. The research examined possible accidents and fatalities involved with each activity, investigated risk factors of such accidents, and finally, identified data needs for safety assessment based on safety regulations and best practices. The automated safety assessment method was then developed using informed and interpreted data for understanding whether a working environment is safe or unsafe.

Accidents in Earthmoving and Surface Mining Activities

As a first step for safety assessment, accident categories of earthmoving and surface mining activities—specifically, loading, hauling, and dumping operations— were investigated. Accidents resulting from maintenance or ingress/egress were not considered; only accidents caused by equipment operation were examined.

According to the literature reviewed (NIOSH 1998, MSHA 2001), the loading operation might cause "rolled overs" (i.e., quarter rolls and other rolls on the same or lower level), "collisions" (i.e., collision with mobile equipment or other large stationary objects), "bounced or jarred" accidents (i.e., a sudden release of energy that caused the machine to bounce or lurch forward or backward), "pinned between" accidents (i.e., pinning between the bucket and frame of skid steer loaders or between the lift arms and frame), and "contacted power line" accidents (i.e., contact with overhead power lines). In addition, the Mine Safety and Heath Administration (MSHA 1999) and the National Institute for Occupational Safety and Health (NIOSH 2001) classified accidents related to hauling operations. The accidents mainly contained "fell over road edge" (i.e., traveling over a road edge and falling down to rest at a lower level), "hung up on road edge" (i.e., traveling onto a road edge and getting stuck without falling over), "rolled overs," "collisions," bounced or jarred" and "contacted power line". MSHA (2001) and NIOSH (2001) also investigated accidents at a dumping site. The common accident types included "fell over the edge" (i.e., traveling through berms and falling over the edge), "hung up on edge," "roll overs," "collisions," bounced or jarred," and "contacted power line."

Risk Factors of Accidents

Following the investigation of accident causes, risk factors contributing to accident potential were analyzed. Figure 1 shows an example of risk assessment diagrams on the dumping operation.

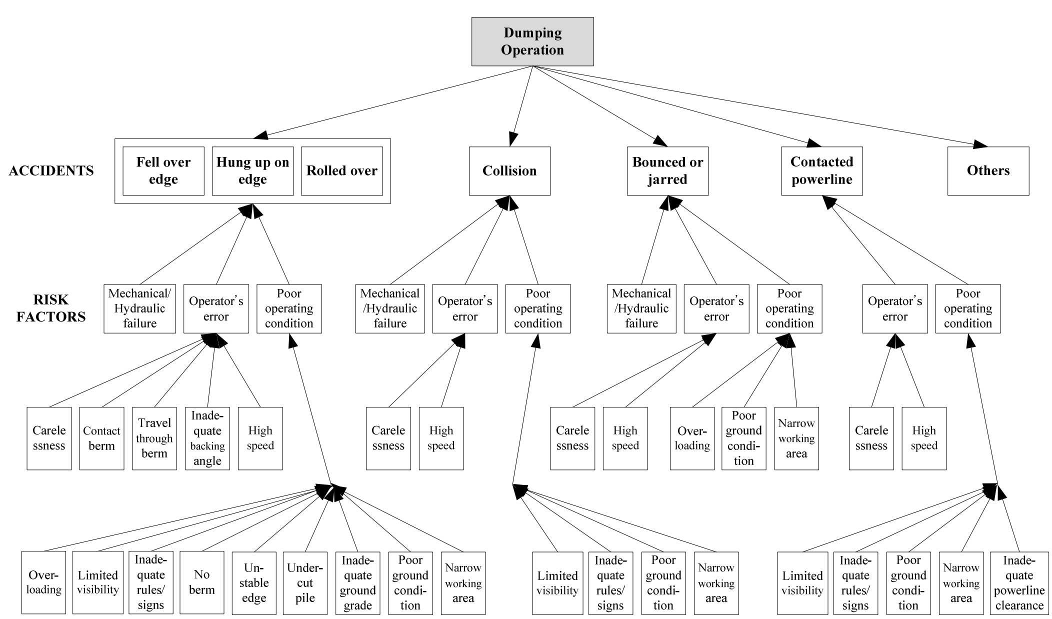


Figure 1 Risk assessment diagram on dumping operation

In general, mechanical or hydraulic failures such as defective brakes and rollover protective structure, careless attitudes of operators, an excessive rate of operation speed, inadequate rules and signs, a congested working area, and poor ground surface conditions such as uneven ground and icy surface conditions can result in any kind of accident (risk factors in Figure 1 - MSHA 1999, NIOSH 2001, MSHA 2001). Poor site layout, a curved road, or large-scale heavy equipment machinery may create limited visibility, and accidents might happen at blind spots with limited visibility. Overloaded material can influence machine rollover, bouncing, or lurching. Power lines that are close enough to the ground can be contacted by operating equipment. Operation-specifically, the undercutting of a material stockpile, that is, removing material from the base of a pile so that it compromises the stability of the pile, may result in instability of edge conditions in the loading and dumping operations, and pile collapse can cause rollover of machinery. A berm has been defined as "a pile or mound of material intended to assist in preventing mobile equipment from traveling over the edge of a bank. Berms are normally used along the edge of haulage roads and dump sites" (NIOSH 2001). Poor berm conditions or missing ones may cause "fell over edge," "hung up on edge," or "rolled over" accidents in hauling and dumping operations.

The safety assessment process proposed in this research deals with risk factors associated with operator errors among the risk factors examined, since the other two categories, poor operating conditions and mechanical/hydraulic failure, lean more toward design and maintenance perspectives. Specific risk factors causing operator errors include careless operation such as berm contact or inadequate backing angle, high operation speed, and traveling through an edge or a dangerous area.

Best Practices and Data Needs for Safety Assessment

The Mine Safety and Health Administration enforces the Mine Act and Title 30 of the Code of Federal Regulation (30 CFR) (MSHA 2008). For every risk factor, meeting best practices in terms of safety regulations were discussed, and finally, data needs to support safety assessment were identified. Table 1 explains best practices mitigating risk factors and identified data needs.

Data needs discussed in Table 1 can be classified into five categories: (1) moving speed of the equipment, (2) stopping distance of the equipment, (3) proximity to strategic spots, (4) proximity to dangerous areas, and (5) proximity to other on-site objects. The strategic spots include a road curve, a hill point, a road intersection, a road edge (berm), and a dumping edge (berm). The dangerous areas contain a specified hazard area, an area between machinery or equipment and the highwall or bank, and an unstable edge of the dumping area. These data can be used as fundamental sources for automated safety assessment of earthmoving and surface mining activities, specifically, loading, hauling, and dumping operations.

Before acquiring identified data, two pre-requisite steps need to be considered. First, "3D object tracking" is necessary because an object's proximity and moving speed can be estimated using three-dimensional information of object positions. Second, "object identification" is also required since safety rules are generally applied differently to different object types. For example, if two haulage trucks are approaching each other, it might be a hazard situation. However, if a loader is approaching a dump truck for material loading, this situation might not be dangerous. In addition, different speed limits need to be applied to different vehicle types. An access authority for the dangerous area can also be assigned only to specific equipment types. For these various reasons, object identification should precede safety assessment.

Pre-requisite Steps: Object Identification and Tracking

In this research, a stereo vision camera was used for acquiring the raw data needed for identification and tracking. Such a camera provides a fast frame rate, feasibility for outdoor applications, long reading range, and the capability for both object localization and 3D modeling. Using this camera, object identification and tracking algorithms were analyzed, modified, and adapted for the proposed construction safety application. Much research has been conducted in the field of computer vision study to develop robust tracking and identification algorithms. The algorithms developed in previous studies (Collins et al. 2001, Stauffer and Grimson 2000, Javed and Shah 2002, Bose and Grimson 2004, Hu et al. 2004) mainly follow these three steps: (1) moving object detection, (2) object correspondence, and (3) object classification. The first step, "moving object detection," aims at separating the regions of motion corresponding to moving objects from the rest of an image (Hu et al. 2004). The second step, "object correspondence," aims at taking the segmented moving regions and matching them to find a corresponding region within an image sequence. The last step, "object classification," aims at classifying moving regions by using common shapes, appearances, or movements (Stauffer and Grimson 2000). Background subtraction algorithms, morphological image processing techniques, connected component algorithms, and different classifiers were reviewed, customized, and employed for this research. The detailed information can be found in another article by the authors (Chi and Caldas, 2008). The identified and tracked object information is now ready to be used to acquire meaningful data for safety assessment.

Table 1 Best practices (MSHA 1999, MSHA 2001) and data needs

No.	Risk factor	Best practice	Data need
1	High speed	Operators should follow the speed limits selected to keep the equipment operating within the capabilities of their braking systems.	Moving speed
		On curves, the speed must be limited to allow adequate traction.	
2	Traveling through an edge	When a vehicle is rounding a curve, cresting a hill, descending a grade, or approaching an intersection, a potentially hazardous condition may exist if the sight distance is less than the estimated stopping distance.	Sight distance (proximity to a curve, a hill, and an intersection), stopping distance
		Berms should give the driver a visual indication of the location of the roadway edge, and the driver should operate the vehicle without contacting berms.	Proximity to a road edge
		Operators should keep a vehicle back from the edge of a slope by a distance equal to at least the width of the berm.	
		Operators should not attempt to dump over the edge of a pile.	Proximity to a dumping edge
		Operators should back up perpendicular to a berm, not at an angle to the dumping edge.	
		Operators should use a berm as a visual indicator only, not rely on it to stop the truck.	
3	Traveling through a dangerous area	The hazard area should be marked with a warning against entry, and, when left unattended, a barrier should be installed to impede unauthorized entry.	Proximity to dangerous areas (a hazard area, an area between machinery and highwall, and an unstable edge)
		Work or travel between machinery or equipment and the highwall or bank should be prohibited.	
		Access to the unstable edge of the dumping area should be restricted.	
4	Careless operation	If vehicles appear to be following one another too closely, the stopping distance should be used for guidance on the distance that should be maintained between the vehicles.	Stopping distance, proximity to other vehicles
		Operators should check for adequate clearance and visibility, especially blind spots, before operation.	Proximity to other on-site objects

Safety Assessment for Earthmoving and Surface Mining Activities

Estimation of proximity to on-site objects

The proposed safety assessment method estimates proximity. It continuously tracks the 3D positions of heavy equipment machinery and workers, and it estimates the distances between objects. The process first assigns a safety margin that should surround heavy equipment machinery and then monitors other objects' proximity as they approach this boundary. The size of any given safety margin can be determined by the stopping distance of the machinery, which is defined as the traveling distance from the instant the operator perceives a hazard and applies the brakes to the instant the machinery completely stops (CSG, Inc. 2008). This time period was calculated with the assumption that operators of average skill can fully stop the machinery within the stopping distance. The stopping distance can be calculated using the following equation (MSHA 2008, CSG, Inc. 2008):

$$d = \frac{{v_0}^2}{2g\mu} + (t_s + t_0) \cdot v_0 \qquad (1)$$

here, v0 is the velocity of the heavy equipment, which can be estimated by considering the different positions of the object's volume centroids within an image sequence. g is gravitational acceleration (9.81 m/s2). μ is the friction coefficient between the tires and the road. The Mine Safety and Health Administration (MSHA) defined typical values for the coefficient of friction between rubber tires and various road surfaces (Table 2) (MSHA 1999). ts is system response time, and MSHA defined this time based on vehicle gross weight (Table 3) in the regulation "57.14101 Brakes" (MSHA 2008). Last, t0 is operator response time, and MSHA determined one second to be the operator response time for those of average skill (MSHA 2008).

Table 2 Coefficient of friction between rubber tires and various road surfaces

Material	Dry	Wet	Material	Dry	Wet
Concrete	0.90	0.60-0.80	Gravel road, firm	0.50-0.80	0.30-0.60
Clay	0.60-0.90	0.10-0.30	Gravel road, loose	0.20-0.40	0.30-0.50
Sand, loose	0.10-0.20	0.10-0.40	Snow, packed	0.10-0.40	0
Quarry pit	0.65	-	Ice	0	0

Table 3 Estimated system response time based on vehicle gross weight

Gross weight (lbs)	1 - 36k	36k - 70k	70k - 140k	140k - 250k	250k - 400k
System response time (sec)	0.5	1.0	1.5	2.0	2.25

After the classification process, the proposed method determines the gross weight of the classified object using a pre-determined database. For instance, if an object is classified as a backhoe, the process finds its weight from the database and assigns it as 25,000lbs. Using this weight, the system response time can be calculated. Figure 2 shows an overall pipeline for safety margin assignment.

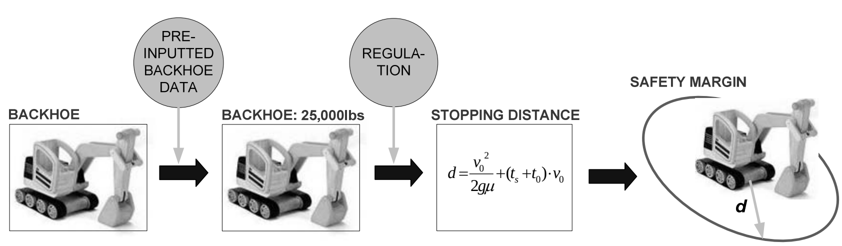


Figure 2 Safety margin assignment

Estimation of proximity to strategic spots or dangerous areas

The proposed method also allows users to define a strategic spot or a danger area within the field of view of the camera. As discussed in the previous section, the strategic spots contain a road curve, a hill point, a road intersection, a road edge (berm), or a dumping edge (berm). The dangerous areas include a specified hazard area, an area between machinery or equipment and the highwall or bank, and an unstable edge of the dumping area. The method first marks such spots and areas. Object tracking and the identification algorithm then keeps monitoring the movement of workers and heavy equipment machinery, and their proximity to the spots or the areas are estimated.

Safety assessment for loading, hauling, and dumping operations

Safety assessment for the loading operation

A loader and a truck are both involved in typical loading operations. The loader scoops material from the stockpile of soil or unformed rock and loads it onto the haulage truck. Since a loading area is generally congested with heavy machinery, different safety rules are applied for different activity types. For example, if two haulage trucks are closely approaching each other, it might be considered a hazard situation. However, if a loader approaches a dump truck for material loading, it might not be dangerous. Because of these differing conditions, travel and working patterns need to be investigated. Figure 3 shows an example of a typical loading zone for surface mining. In Figure 3, the area near the highwall is regarded as a dangerous working area. The proposed safety assessment method continuously tracks the movement of heavy machinery and estimates their proximity to other machinery and pre-determined dangerous areas to facilitate safety decision-making. The actual loading operation, from the instant that the truck stops for loading to the instant that the truck starts hauling away, is considered to be a safe working condition.

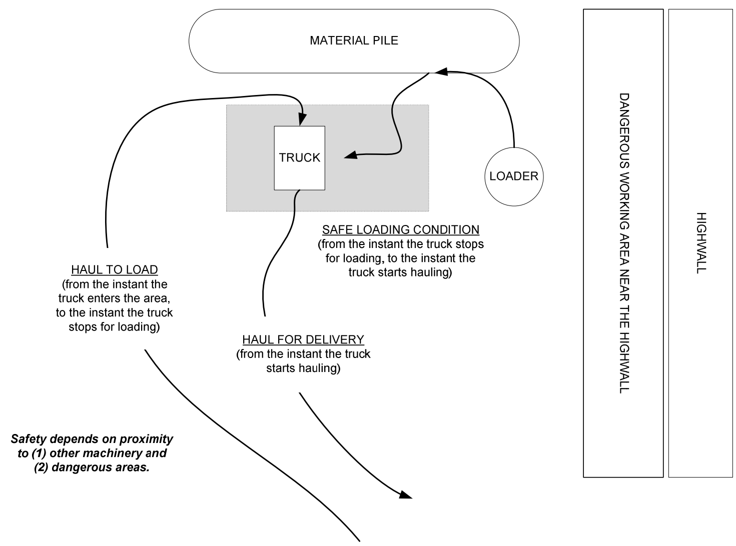


Figure 3 Safety assessments for loading operation

Safety assessment for the hauling operation

The proposed assessment method for the hauling operation first tracks the machine's moving speed, which is one of the most common risk factors of haulage-related accidents. As shown in Figure 4, the method first determines dangerous access spots near the road edge, tracks proximity to these spots, and prevents the truck from traveling through the spots. In addition, the method sets a strategic spot near a road corner, a hill, or an intersection and calculates proximity to the spot in order to help operators have a clear sight distance. The method also estimates the proximity to other trucks and compares it with the calculated stopping distance for safety decision-making.

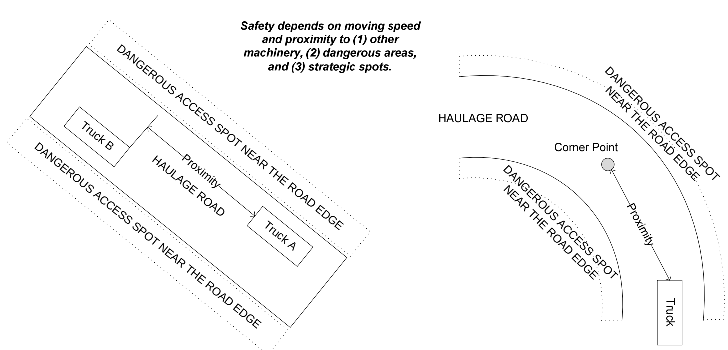


Figure 4 Safety assessment for hauling operation

Safety assessment for the dumping operation

The most common fatal dump-point accidents involve trucks going over the edges of piles. Thus, the proposed assessment method for the dumping operation primarily focuses on the estimation of proximity to the berm near the pile edge (Figure 5). While the dump truck is backing up to the edge, the method estimates its proximity to the berm in order to prevent the truck from contacting the berm and potentially falling over it. The method also monitors proximity to other trucks to avoid collision between machinery.

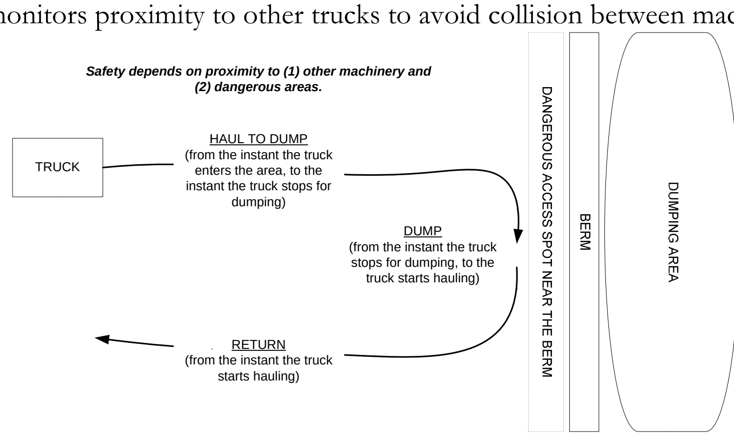


Figure 5 Safety assessment for dumping operation

Preliminary Results

Preliminary experiments were conducted for testing and validating the proposed methods. The algorithm codes were written using the C++ programming language. A laptop computer (3.2 GHz Intel Pentium 4 CPU and 1.5 GB of RAM) was used for program implementation. For analyzing the performance of object identification and tracking, experiments were conducted on an actual construction site, where simple earthworks were performed by a skid steer loader, a backhoe, and a worker. A background subtraction algorithm first extracted the moving objects from an image sequence, and the spatial characteristics of the moving objects were then entered into classifiers (a normal Bayes classifier or a neural network). Using these entered variables, the classifier classified each object as a worker, a loader, or a backhoe. Figure 6 shows examples of the classification results. A total of 1,282 images were analyzed, and the classification algorithm was processed three times per second, which was a close approximation of real-time applications. The overall classification errors of two of the classifiers were under 4% (Bayes classifier: 3.35% and neural network: 3.90%), which showed an acceptable rate of accuracy for the algorithms.

Figure 6 Object identification and tracking: worker, loader, and backhoe

For performance evaluation of the proximity estimation method, outdoor experiments were conducted in a controlled environment using a skid steer loader and a worker at the Field Systems and Construction Automation Laboratory at the University of Texas at Austin. The safety margin around the classified loader was set based on its moving velocity. The proximity estimation process then observed the loader's proximity to the worker when he approached the safety margin determined by the machine's stopping distance. Table 4 shows the estimated proximity between the loader and the worker (distance between two centroids – 3D worker width/2 – 3D loader width/2) and the calculated stopping distance. Since the worker first approached the moving loader and then moved away from the loader, the proximity between two objects decreased first and increased again. When the proximity was smaller than the stopping distance as shown in Table 4, we might say the worker was within the dangerous distance from the operating loader.

For proximity estimation to dangerous areas, the restricted area was manually pre-determined, and the proximity estimation method observed access area violation to when tracked object approached the area. All actual violations were identified. The method determined that the worker crossed into the danger area 22 times over the course of the overall 50 captured images, and the loader entered it 13 times over 41 images.

Table 4 Performance analysis of proximity estimation

Frame	Distance b/w centroids(m)	Worker width (m)	Loader width (m)	Proximity (m)	Loader vel. (m/s)	Stopping distance (m)
1	3.70	1.46	4.39	0.77	0.50	0.79
2	3.02	1.32	4.70	0.11	0.48	0.77
3	2.90	1.04	4.65	0.06	0.46	0.72
4	3.16	1.82	4.05	0.23	0.47	0.74
5	3.70	1.69	4.32	0.70	0.51	0.80

Conclusions

This paper presents an ongoing research project concerning the development of an automated safety assessment framework for earthmoving and surface mining activities. Preliminary results showed the feasibility of proximity estimation as well as object identification and tracking, which can be used for automated safety assessment in future research. In order to evaluate the performance of the entire safety assessment process, experiments on an actual surface mining project are in the planning stages. Dangerous areas and strategic spots will be first identified and assigned at loading, hauling, and dumping sites, and the proximity to these areas and spots will then be monitored in 3D for safety decision-making. The proximity of on-site objects to each other will also be estimated, and these distances will be compared with the stopping distances of the machinery. Last, information of static objects on sites can be pre-inputted into the process so as to achieve more robust safety decisions.

References

[1] Ahmad, K., and Gibb, A. (2004). “Towards effective safety performance measurement – evaluation of existing techniques and proposals for the future.” Construction Management Systems, Spon Press, London, UK.

[2] BLS (Bureau of Labor Statistics). (2003, 2004, 2005, 2006). “Census of fatal occupational injuries.” Dept. of Labor, Washington, DC. <http://stats.bls.gov/iif/oshcfoi1.htm> (June. 11, 2008).

[3] Bose, B., and Grimson, E. (2004). “Improving object classification in far-field video.” Proc., IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Washington, DC.

[4] Chi, S., and Caldas, C.H. “Automated object identification using video cameras on construction sites.” Submitted to the Journal of Computer-Aided Civil and Infrastructure Engineering, Special Issue on Advances in Construction Automation (In Review).

[5] Chua, D.K.H., and Goh, Y.M. (2004). “Incident causation model for improving feedback of safety knowledge.” Journal of Construction Engineering and Management, 130(4), 542-551.

[6] Collins, R.T., Lipton, A.J., Fujiyoshi, H., and Kanade, T. (2001). “Algorithms for cooperative multisensor surveillance.” Proc., IEEE, 89(10), 1456 - 1477.

[7] CSG (Computer Supporting Group), Inc. (2008). “Vehicle stopping distance calculator.” <http://www.csgnetwork.com/stopdistcalc.html> (March. 3, 2008).

[8] Hu, W., Tan, T., Wang, L., and Maybank, S. (2004). “A survey on visual surveillance of object motion and behaviors.” IEEE Trans. on Systems, Man, and Cybernetics, 34(3), 334-352.

[9] Javed, O., and Shah, M. (2002). “Tracking and object classification for automated surveillance.” Proc., 7th European Conference on Computer Vision (ECCV), Copenhagen, Denmark.

[10] MSHA (Mine Safety and Health Administration). (1999). “Haul road inspection handbook.” PH99-I-4, Dept. of Labor, Washington, DC.

[11] MSHA (Mine Safety and Health Administration). (2001). “Dump-point inspection handbook.” PH01-I-6, Dept. of Labor, Washington, DC.

[12] MSHA (Mine Safety and Health Administration). (2008). "Title 30 CFR." Dept. of Labor, Washington, DC. <http://www.msha.gov/30CFR/CFRINTRO.HTM> (December. 8, 2008).
[13] Navon, R., and Kolton, O. (2007). "Algorithms for automated monitoring and control of fall hazards." Journal of Computing in Civil Engineering, 21(1), 21-28.
[14] NIOSH (National Institute for Occupational Safety and Health). (1998). "Preventing injuries and deaths from skid steer loaders." 98-117, Dept. of Health and Human Services, Washington, DC.
[15] NIOSH (National Institute for Occupational Safety and Health). (2001). "Haulage truck dump site safety: an examination of reported injuries." IC 9454, Dept. of Health and Human Services, Washington, DC.
[16] Stanton, W.A., and Willenbrock, J.H. (1990). "Conceptual framework for computer-based construction safety control." Journal of Construction Engineering and Management, 116(3), 383-398.
[17] Stauffer, C., and Grimson, E. (2000). "Learning patterns of activity using real-time tracking." IEEE Transactions on Pattern Recognition and Machine Intelligence (TPAMI), 22(8), 747-757.

A Proactive System for Real-time Safety Management in Construction Sites

Alessandro Carbonari[1], Berardo Naticchia[2], Alberto Giretti[2] and Mario De Grassi[2]

[1]Polytechnic University of Marche, DACS Department, Engineering Faculty, Ancona, Italy. Tel: +390712204397; Fax: +39 0712204582; Mobile: +39 3402502768, alessandro.carbonari@univpm.it
[2]Polytechnic University of Marche, Smart Space Solutions SRL, Via Brecce Bianche, 60131 Ancona, Italy

Abstract

The purpose of this paper is presenting a new advanced hardware/software system, boasting two main features: first it performs real time tracking of workers' routes in construction sites; then it implements an algorithm for preventing workers to be involved in hazardous situations. This research step is part of a wider ongoing research concerning the development of a new generation of advanced construction management systems, allowing for real-time monitoring and coordination of tasks, automatic health and safety management, on-site delivering of technical information, capture of as-built documentation. Exploiting the high accuracy provided by the UWB system responsible for position tracking and successfully tested in previous research, our software interface is able to graphically reproduce (and store) the travel patterns of workers. Moreover, it constantly checks if they are accessing hazardous areas, using an algorithm based on a predictive approach: it is conceived to predict in advance whether any worker is approaching a forbidden area, in fact performing virtual fencing. This approach could be easily extended to other applications, too. Some preliminary tests simulated in the DACS laboratory are described and the obtained results discussed.

Keywords: UWB tracking system; health and safety in construction; virtual fencing; automated monitoring.

1. Introduction

It is well accepted that construction sites' management is a very tough task, due to a number of reasons (Naticchia et al. 2005): they are custom designed, they host the realization of very complex building projects, involving thousands of parts and components, and changes of design plans at construction time are not uncommon. Hence, human resource and facility management would need a continuous monitoring, in order to be correctly performed. In addition, their organization is very strongly related with health and safety objectives. Considering that coordinators for health and safety matters at the construction level cannot always attend construction works while they are ongoing, it can be inferred that the practical application of the prescriptions required by the health and safety plan (according to European Council Directive 92/57/EEC) are left to workers. In addition, also coordinator's enduring presence on construction sites would not be adequate to prevent accidents in very large sites, as not capable of ubiquitous control. This situation often leads to scarce application of health and safety requirements, that could be the cause of dramatic accidents.

However, a set of recently marketed technologies could provide the necessary background for developing a new generation of real-time construction management systems, which can be seamlessly integrated into the actual arrangement of the construction work (Abderrahim et al. 2005). The two main kinds of technologies are: on one side, accurate tracking technologies, that could provide in real time the position of all workers present in a construction site; on the other side, tools for behaviour capture, used to model and then predict assets' behaviour, in order to infer what is going to happen.

The good aspect of tracking technologies is that some of them are very accurate, and are able to track workers by just equipping them with such small embedded tags, that they do not interfere with ongoing activities. It obviously needs a system setup to cover all the area of interest. Considering that they will give back in real time the position of every employed resource, then it could be interfaced with a system having the task of predicting what is going to happen and analyze if it could trigger some dangerous situations. It would have the advantage of improving presently adopted approaches for safety management and optimizing resource usage and working efficiency at the same time. The tools developed for behavioural model capture would have the task of discerning dangerous situations in advance.

This paper concerns a feasibility study of a pro-active system for health and safety management in construction sites, fully described in paragraph 3, and mainly composed by two parts: the first performing real-time position tracking, while the second providing real-time prediction of risky events. The results obtained from preliminary experimental tests performed in our laboratories will be described and discussed to plan further steps.

2. State of the Art

Despite mobile computing for production management has already been developed and reliably applied to manufacturing (Fuller et al. 2002, Brewer et al. 1999), it is almost completely missing in construction sites, due to its outdoor, heterogeneous and highly evolving nature. Some preliminary results have been obtained in construction sites through the application of standard outdoor position tracking technologies. Some researchers have developed a mechatronic helmet, equipped with a GPS antenna and a bidirectional communication system for workers' safety control (Abderrahim et al. 2005). A lot of efforts have been made to explain the main advantages that would be pursued with the adoption of mobile computing (Rebolj et al. 2001). Among them, the most important ones, devoted to the field of construction management, are: supply delivery records and progress updates directly at the jobsite; rapid communication and collaboration throughout the entire project life cycle, from financing and planning through engineering and design, procurement, construction, and facility management; systems that provide construction teams with a project-specific extranet, whereby remote team members can access up-to-date documents.

For that reasons, the main efforts have been directed towards mobile computing technologies development and resource tracking. Within the first field of research there are some very valuable works: an integrated RFID and GPS technology for the purpose of tracking highly customized prefabricated components and avoiding delays in construction (Ergen et al. 2007); embedding RFID tags in building components to store design data, which can be passed to the people in charge of maintenance during its operational phase (Cheng et al. 2007); efficiency improvement of tool tracking and availability increase by using RFID tags (Gajamani et al. 2007).

On the other side, there have been several attempts to develop systems for continuous resource monitoring, adopting a wide range of technologies, from Rfid to GPS. However, the only one that was shown to be really capable of providing good performances for very complex places, resulted to be Ultra Wide Band: fieldwork tracking in an iron made construction showed a position accuracy significantly better than 1 m (Teizer et al. 2007), the same happened for reinforced concrete made construction buildings (Giretti et al. 2008). It suggests that this technology would be very suitable for safety management, were accurate localization is critical to avoid workers' access into hazardous areas, to perform collision avoidance, to prevent falls from height and a number of accidents. Other research is facing the problem from a higher perspective: for example, a Project called "SightSafety" has the main aim of developing innovative and proactive systems for automated collision detection avoidance, which is part of a more ambitious project regarding health and safety management in constructions sites (Riaz et al. 2006). Very interesting is the two tier software system the authors have developed, whose business logic has been designed to recognize whether dangerous situations take place. It is also stated that the higher the position accuracy, the better the system performs its tasks. That remarks the importance of using some very accurate location systems like UWB.

Therefore, now big interest is focused on how tracking systems could be interfaced with software systems to recognize dangerous situations and which kind of algorithms should be developed to pursue this last objective. Among all the tested approaches, there is one attempt consisting in recording time series paths (derived from UWB tracking), in order to identify those ones which are safer by the frequency workers pass across them and then signal when some workers would choose a not safe path (Teizer et al. 2008). This approach is very similar to what proposed in other fields of research, where algorithms for behavioural models are developed: they were used to forecast workers' behaviour, in order to perform risk avoidance in advance with respect to the time when risky situations are experienced (Wang et al. 2004): thus, alert messages may precede dangers. Our paper rises from such considerations and tries to propose a first software application for virtual fencing of dangerous areas in construction. It laid the basis for further development about the production of the final system release, to be installed and tested in real construction sites.

3. The New Proactive System

3.1 Overall Description

Figure 1-a provides a scheme of the system under development, that will be made up of some hardware sub-systems, interfaced with a software application for whole management. It is expected that tracking very complex and large construction sites will be performed by installing several UWB sub-networks, connected to one service application, implementing a software tool for data and logic management. In this first period of research we are focusing on health and safety, while other features will be added later (e.g. providing "on-demand" information to personnel operating on site, automatically producing "as-built" documentation, performing real-time management even through remote control), and exploiting the same UWB localization system. This application software is thought to implement a three tier software architecture for real-time construction management (Figure 1-b). The lowest level implements sensor, localization, communication, inference and data management logics. The middle level "business logics" implements high level task oriented functionalities (e.g. when and how managing virtual fencing or collision avoidance tools, included in the lowest level, transferring those data to the high level application modules, according to users' requests). The application layer (highest level) customizes general business logic functionalities to specific application domain, such as health and safety management. In this paper we will focus first on development of the inference API, and then on the relationship between it and the position tracking module, which is part of the localization API.

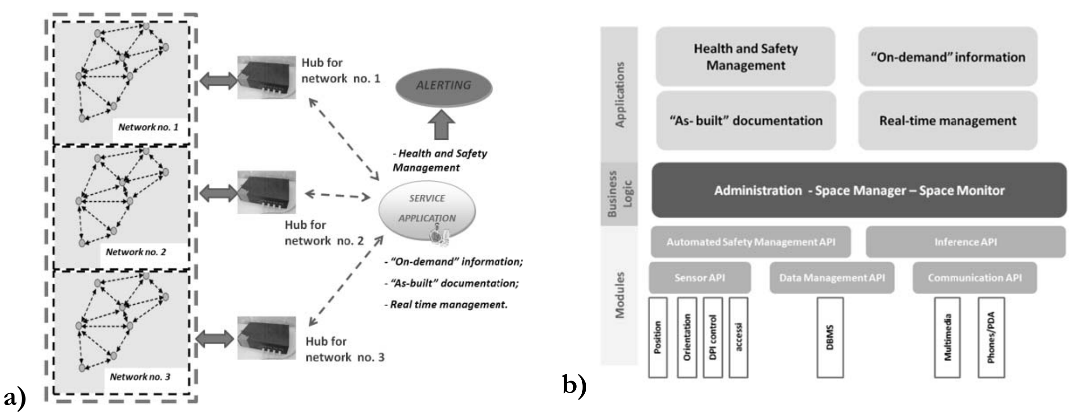


Figure 1. Schematization of the pro-active system logic (a) and three-tier software system implemented in the Service Application (b).

Then, a preliminary user graphic interface for the visualization of the tracked tags will be presented. For the purpose of this preliminary study, it was decided to develop a tool for the preventing workers access into previously defined dangerous areas. The algorithm is fully described in paragraph 3.3, while the graphic interface for input insertion and visualization is shown in Figure 2-a. It is subdivided into six areas: tag-user configuration, accepting user inputs and storing correspondence with their IDs; uploading map image; size and map insertion in the reference system of the monitored area; not accessible area configurations, that requires the insertion of the coordinates of the area to be not accessed; map visualization; connection to the UWB tracking system (defined in the same geographic system).

3.2 System's Expected Performances and Previous Research

Previous experimental tests led in the construction site shown in Figure 2-b and 2-c, let us conclude that UWB tracking systems can perform very accurate location tracking at least until reinforced concrete structure's erection (Giretti et al. 2008). Using just four receivers and equipping every resource with 1 W tags, it was possible to monitor a 500 m2 large construction area. Hence UWB overcomes the limits of GPS, which was not able to track indoor.

Instead, a different and more intense configuration for UWB receivers will be necessary to perform monitoring when masonry buildings are monitored. In any case, it came out that UWB tracking is very good for tracking from the beginning of the construction progress, until structure frame is built. Hence it would be able to support several automated features (e.g. collision avoidance, virtual fencing).

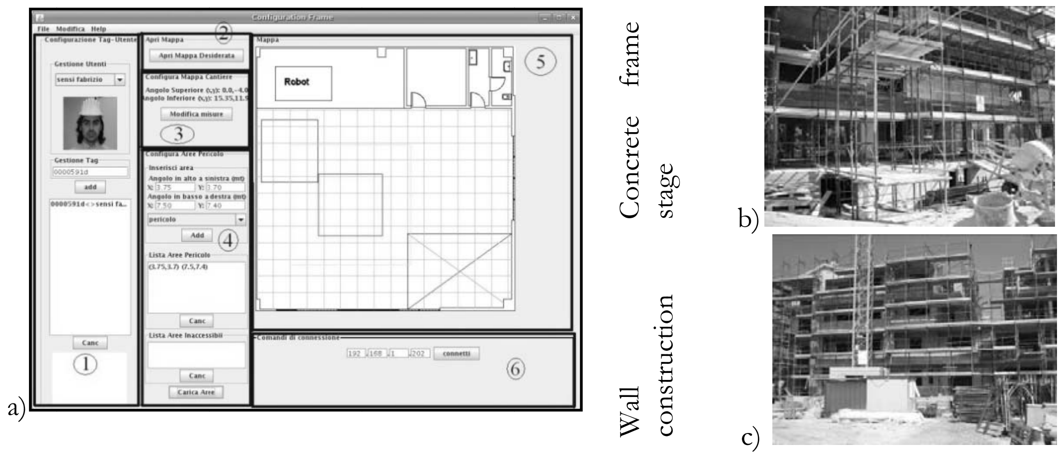


Figure 2. Graphic interface of the software system (a) and the two progress stages of the construction site used for UWB testing (b, c).

3.3 Predictive Approach to Risk Management

It is generally stated that, when coping with risky situations, it is necessary to adopt a predictive approach (Wang et al. 2004). An intelligent approach is always necessary to check risky events in advance and render warning signals with suitable timing for workers, in order to prevent bad consequences. In addition, the high polling frequency of UWB technology (from 1 to 60 Hz) should be exploited to process real-time data and properly manage those situations by providing alerting signals in real-time. We think that two basic intelligent approaches are available to manage those situations, whose classification is valid also in the particular case of virtual fencing of dangerous areas, considered for the purposes of this paper:

1. a deterministic approach, that predicts dangerous situations starting from the actual route stepped by workers;
2. a probabilistic approach, which could be based on probabilistic models (e.g. Bayesian Networks), to perform knowledge capture first and inference processing later, in order to predict first the route that every monitored worker is going to follow and then whether this path is expected to lead towards risky situations (Howden et al. 2003).

In this paper, we have chosen to analyze the first approach, even if the second one will be studied soon later, within the next research step. Hence, we are proposing a first predictive method to manage risks in real-time.

The software application interfaced with the UWB tracking system implements an algorithm for virtual fencing. Its main objective is checking risks in advance, in order to warning endangered operators before they could be harmed. For that purpose, the first step is discerning and localizing where dangerous areas are placed. They could be those areas were presence of workers is not allowed (e.g. close to scaffolding not equipped with protections against falling), which may be selected directly from the site layout, during the design stage. There could also be the possibility those areas are dynamically changing: in case collision avoidance between two assets is performed, one object could be processed as the area not to be entered and the other one as the operator moving throughout the construction site. Hence this first step is critical, being the basic tool for further applications. The main assumptions made for this first technologic development are:

- the dangerous area is static and its location known in advance;
- risks must be predicted before they take place.

The general approach followed in order to pursue this objective consists in surrounding every forbidden area by a "warning" strip (in our drawings, the dangerous area is usually red coloured, while the warning strip is yellow, like in the following Figure 5-a). Warning strip's thickness must be sized according to tag frequency and to average operators' speed. This second parameter was assumed to be fixed at 0.5 m/s, while the first was sized through adequate testing (paragraph 4.1).

In this way, the algorithm is able to check the behaviour of operators who have already entered the yellow area, in order to understand whether they are approaching dangers. Figure 3 better explains our algorithm's logic. It first checks whether any operator has already entered the red dangerous area: in the positive case a red alarm is sent, otherwise no alarms at all are sent and the algorithm checks if the operator has entered the yellow surrounding strip. If this is true, then its present position is compared with the previous one (recorded in the previous time step) and a warning alarm is sent in case he has got closer to the border line; in the opposite case the whole procedure is repeated. It is straightforward that tracking operators' movement within the yellow strip is critical for this application.

Figure 3. The algorithm's logic

4. Experimental Campaign

4.1 Preliminary tests

The experimental campaign was aimed at validating the algorithm proposed and evaluating whether the UWB system adopted is enough for the application we are suggesting in the field of health and safety management. All the tests have been performed in the laboratories of the Department of Architecture Construction and Structures (DACS) of the Polytechnic University of Marche. They have been planned in such a way to be rather representative of the actual situations expected in construction sites. Two experimental phases have been carried out:

- the first one to validate the simplified version of our algorithm, and to infer how thick should the yellow warning strip be around the red area;
- the second one to validate the final release of the proposed algorithm and conclude about its reliability.

The first set of experiments simulated the presence of a not accessible area in a construction site, and wanted to implement a reduced version of the control algorithm, made up of just the first part, signalling when a worker accessed the dangerous area. For that reasons, we measured the gap between the time when a resource crosses the border line and the time when the software alerts about that. Four UWB receivers have been placed far from the corners of one (3.7x4) m large red area (Figure 4-b). Our hardware/software system, interfacing the UWB system's hub to our software application tool (developed in JavaTM environment) and implementing the algorithm of Figure 3, was setup. Then one of our collaborators was equipped with a 1 Hz tag and was asked to enter repeatedly the red area following the three paths of Figure 4-a. Measures were repeated 8 times for each path and it came out that the application tool is able to send warning signals (Figure 4-c) within a maximum distance of 1.2 m beyond the border line, with an average distance gap of 0.7 m from it. Hence a warning strip slightly wider than 1.5 m should be planned around the red area with 1 Hz tags, in order to face also rare events (e.g. some tag bursts are missing to more than one receiver and position is computed in the next iteration, in fact postponing alert signals).

Figure 4-c shows one of the red alarm signals sent by the application software after the boy simulating a worker passed beyond the border line (Figure 4-d). Our system worked very well at this stage, with no failures, being the gap determined just by the time required by the UWB system to update tags' location, which is inversely proportional to its frequency. Hence, in the following test, a 1.7 m thick warning strip around the red area was drawn on the laboratory's floor (Figure 5-a), simulating the strip inside which the

algorithm checks resource's behaviour before sending a warning signal, which depends on his intention to get closer to the border line.

a b **c** **d**

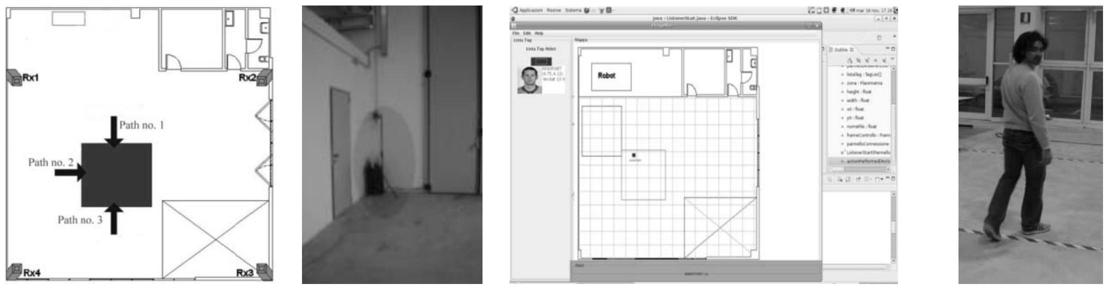


Figure 4. Red area position and the three paths used for the 1st experimental phase (a);one receiver (b); screenshot of the application sending a red alarm signal (c, d).

4.3 Validation of the algorithm's logic

After having arranged our laboratory as depicted in Figure 5-a, one of our colleagues was asked to step the three paths shown in figure, in order to collect 10 measures for each of them and check for errors. Three types of potential errors for the system have been preliminarily discerned:

- absence of alarms: the system does not send alarms, even if it should do.
- false red alarm: the system send red alarm even if the worker is again in the warning strip;
- false warning alarms: the system send a warning alarm even if the worker is not getting closer to the border line.

One different set of tests has been performed for each path in Figure 5-a. For path no. 1, no false alarms have been collected. For path no. 2 just one error of type 1 has been collected (due to the low tag rate of 1 Hz, that did not allow to monitor the worker twice before accessing the red area). For path no. 3 three errors of type 1 have been collected, which were then shown to be due to interfering objects very close to receiver no. 2. Figure 5-b shows the graphic interface sending a warning "yellow" alarm, due to a worker approaching the border line.

From those results it was concluded that the algorithm's logic is correct and works properly, just in case location measures are accurate. The different performances observed among the several measures was dependent just on the random variation of location measures, proper of the UWB system. In order to better explain that concept, the worst case for the system was experimented: as in Fig 5-c, one of our collaborators was asked to walk parallel to the border line. It came out that every measure was affected by one or two errors of type 3, due to the standard localization accuracy of UWB system, whose measures are randomly varying around the real value within a 0.3 m distance. Due to that, the system could send false alarms. In order to reduce that possibility, two approaches could be followed:

- increasing tags' frequency, hence reducing the yellow strip's thickness and the probability that a worker could walk parallel to the border line;
- reducing random variation around the real value, through the use of statistically filtered tracking measures.

5. Conclusions and Future Research

In the preliminary analyses carried out in this paper, the performances of a virtual fencing application tool have been simulated. Tests performed in our laboratories showed that our deterministic algorithm is rather reliable to perform virtual fencing of dangerous areas, using a predictive approach. Further amendments will be necessary to face also dynamic and evolving environments like construction sites, that will require to integrate our tool within a higher software level for behavioural modelling of those environments. In any case this tool is almost ready to be applied in static contexts where collision avoidance between resources must be performed.

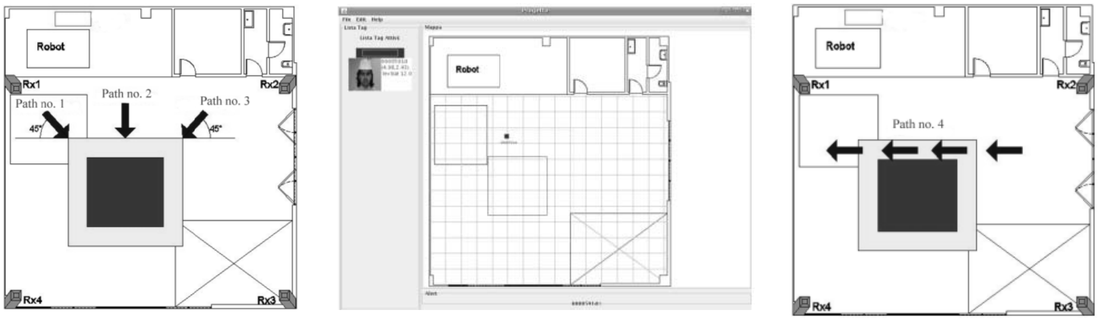


Figure 5. Laboratory setup for testing algorithm's reliability (a), screenshot of the spatial reasoning API's virtual fencing tool (b) and path no. 4 (c).

The few false alarms encountered could be solved by adopting two amendments: filtering localization data, that will minimize the random variation admitted by UWB system around the real position; using tags with a higher frequency, that would allow to decrease the warning strip's width and increase the algorithm's reliability.

Finally, other probabilistic techniques will be developed, to strengthen predictive capabilities of this application.

References

[1] Naticchia, B., Giretti, A. & Giretti, P. (2005) Systems for Real Time Construction Site Management, Proceedings of the Joint 2006 CIB Symposium W065/W055/W086, Rome, Italy, October 2006.

[2] Abderrahim, M., Garcia, E.; Diez, R., Balaguer, C. (2005) A mechatronics security system for the construction site, Automation in Construction, Vol. 14, No. 4, 460-466.

[3] Fuller, S., Ding, Z., Sattineni, A. (2002) Case Study: Using the Wearable Computer in the Construction Industry, Proceedings of the 19th ISARC, National Institute of Standards and Technology, Gaithersburg, Maryland, September 23-25, 551-556.

[4] Brewer, A., Sloan, N., Landers, T. L. (1999) Intelligent tracking in manufacturing, Journal of Intelligent Manufacturing, Vol. 10, No. 3-4.

[5] Rebolj, D., Magdič, A., Čuš Babič, N. (2001) Mobile computing in construction. Advances in Concurrent Engineering, Proceedings of the 8th ISPE International Conference on Concurrent Engineering - Research and Applications, West Coast Anaheim Hotel, California, USA, July 28 - August 1.

[6] Ergen, E., Akinci, B., Sacks, R. (2007) Tracking and locating components in a precast storage yard utilizing radio frequency identification technology and GPS, Automation in Construction, Vol. 16, No. 3, 354-367.

[7] Cheng, M. Y.; Lien, L. C.; Tsai, M. H.; Chen, W. N. (2007) Open-building maintenance management using RFID technology, Proceedings of the 24th International Symposium on Automation and Robotics in Construction – ISARC 2007, September 19-21, Kochi, India.

[8] Gajamani, G. K., Varghese, K. (2007) Automated project schedule and inventory monitoring using RFID, Proceedings of the 24th International Symposium on Automation and Robotics in Construction – ISARC 2007, September 19-21, Kochi, India.

[9] Teizer, J.; Lao, D.; Sofer, M. (2008) Rapid automated monitoring of construction site activities using Ultra-WideBand, Proceedings of the 24th International Symposium on Automation and Robotics in Construction – ISARC 2007, September 19-21, Kochi, India.

[10] Giretti, A., Carbonari, A., Naticchia, B., De Grassi, M. (2008) Advanced real-time safety construction management system for construction sites, Proceedings of the 25th International Symposium on Automation and Robotics in construction, Vilnius, Lithuania, June 26-29.

[11] Riaz, Z., Edwards, D.J., Thorpe, A. (2006) SightSafety: a hybrid information and communication technology system for reducing vehicle/pedestrian collisions, Automation in Construction, Vol. 15, No. 6, 719-728.

[12] Teizer, J., Mantripragada, U., Venugopal, M. (2008) Analyzing the travel patterns of construction workers, Proceedings of the 25th International Symposium on Automation and Robotics in Construction – ISARC 2008, June 26-29, Vilnius, Lithuane.
[13] Wang, A. P., Chen, J. C., Hsu, P. L. (2004) Intelligent CAN-based Automotive Collision Avoidance warning system, Proceedings of the 2004 IEEE International Conference on Networking, Sensing & Control, March 21-23, Taipei, Taiwan.
[14] Howden, N., Curmi, J., Heinze, C., Goss, S., and Murphy, G. (2003) Operational Knowledge Representation: Behaviour Capture, Modelling and Verification, Proceedings of the Eighth International Conference on Simulation Technology and Training (SimTecT '03), Adelaide Australia.

Development of Warning System for Preventing Collision Accident on Construction Site

Soungho Chae

Kajima Technical Research Institute, Tobitakyuu2-19-1, Chohushi, Tokyo, Japan, chae@kajima.com

Abstract

This paper describes the development of a personal warning system for preventing construction site accidents involving workers and heavy equipment. The required functions were established on the basis of a past accidents model. The system was designed to send a warning via wireless from a server computer to a worker and relevant parties according to estimation of distance between worker and equipment. RFID technology was used to estimate the distance. A warning message was defined as a means of sending information about danger to the worker and relevant parties. Parties receiving a warning were set up for each case, and the contents of the warning message were prepared accordingly. An experimental prototype device for sending a message from a server computer to a smart-phone was made under a wireless local area network, and it was clarified that to accept a warning message and to send a response.

Keywords: Safety management, Collision accident, Warning, RFID, Wireless device

Introduction

Several types of vehicles and heavy equipment, such as cranes and excavators, are used on construction sites to increase labour efficiency. This kind of equipment plays a crucial role in planning the work processes of building structures. The construction manager is required to plan vehicle zones and safety facilities. However, the number of accidents in which workers are struck by heavy equipment caused by negligence is increasing. In 2007, in the Japanese construction industry, more than 3,800 persons (about 19% of the total) suffered injuries due to collisions with vehicles or heavy equipment that required 4 days or more for recovery [1].

Construction managers recognize the need to address this problem through a comprehensive approach to accident prevention. A method is necessary to complement accident prevention techniques such as Toolbox Safety Talk and Hazard Identification, which are conventionally performed to assist in avoiding accidents. The author has designed a Collision Accident Prevention System (CAPS) that uses an RFID device based on accident issues [2].

The objective of this study was to develop a warning system as a part of the CAPS. Past collision accidents were analyzed to clarify causative factors. The basic function of the warning system was thus designed on the assumption that data on the location of heavy equipment and workers is available. Finally, a concept of a warning device was proposed with a user interface and communication protocol. Function of Waring system

Basic Concept

CAPS supports accident prevention involving vehicles, heavy equipment and workers using working area data of each object. The working area is the space bounded by the reach of heavy equipment. An RFID (Radio Frequency Identification) device and a data communication system are used to estimate the size of the working area. The CAPS monitors the working area of all objects on the site and calculates the positional relation of heavy equipment and workers. When it is determined that a worker is in the restricted area of heavy equipment, a message warning of the danger of a collision is sent to the worker and concerned parties.

Analysis of Past Accidents

Past accidents involving heavy equipment are analysed to model error occurrences that could not be prevented by common warning tools using, e.g., light or sound. The error model is made for hydraulic excavators, lifts, truck cranes, tower cranes, etc. The items of the model are as follows.

1) Contact part of equipment

2) Accident location
3) Task of worker
4) Error of worker, concerned parties, or manager causing accident

For example, the model of accidents involving hydraulic excavators shows the bucket as the most common contact part and the body as the second. Accident location was below or at ground level, where excavator usage is assumed, and the worker's task was work control, ground arrangement, excavation assistance, etc. Workers' errors included entering within the operating range of the machine, assuming a forced work posture, and wearing insufficient safety gear. Errors of concerned parities, i.e., excavator operators in this model, were lack of awareness of a worker's presence before starting to operate the machine. Managers' errors included defective safety plans, lack of a spotter, and ignoring of safety training. The modelling result clarified the factors affecting collision accidents.

Setup of System Functions

The author set up the following five functions for a warning system on the basis of the past accidents model.

1) Send worker a warning

This function reduces instances of workers ignoring warnings, which is likely to occur in conditions of continuous unrelated warning beeps. To enable him to recognize a danger, a warning is sent to a worker who could potentially access a heavy equipment working area. A direct warning message to a worker is expected to raise his awareness of a potential collision.

2) Send concerned parties information

This function enhances collision avoidance behaviour of a worker who received a warning. Concerned parties become aware of the collision potential and direct the worker to a safe area. Concerned parties include operators, spotters and signal men.

3) Ensure time for awareness

This function provides sufficient time for a worker who has received a warning to avoid a dangerous situation. A prompt warning enables the worker to escape from danger without panic or hesitation and the operator to check the movement range of his machine.

4) Set warning contents according to task

This function prevents workers from adopting a blaze attitude due to warning habituation. A worker who receives the same warning for all situations could try to avert the danger without confirming the true situation. The worker is required to confirm the dangerous situation according to the particular item of equipment and situation based on the warning information.

5) Require worker to respond

This function supports the judgment of whether a manager should directly warn a worker who is in a dangerous area. A reply is required of a worker who has received a warning as the result of recognition and avoidance of danger. At the same time, a list of workers who disregard the warning is sent to the concerned parties and the safety of worker is verified by the concerned parties.

Warning System Design

Area Definition for Warning

The author defined the working area of heavy equipment as a danger zone, and set up the area surrounding the danger zone as a confluence zone and safety zone [Figure-1]. The size of danger zone is determined by maximum range of equipment's movement. The confluence zone is designed according to the migration speed of equipment, the shape of site and the process of project. A worker who is in the confluence zone is at risk of becoming in a dangerous situation. CAPS was designed to send a warning to a worker who is monitored to be in a confluence zone or a danger zone.

System Flow

CAPS sends a warning by wireless from the administrative server to a worker and to relevant parties according to the estimation of distance between workers and equipment [Figure-2]. If a worker is in the confluence zone, his information is collected from a database and necessary cautions are sent to him and to relevant persons such as operators and spotters.

A reply to a warning is required in order to verify reception and execution. If there is no reply within a specified time from a worker who is in the confluence zone, the relevant persons are required to confirm the situation and to ensure the worker's safety.

If a worker is in the danger zone, although emergency stop and alarm activation are necessary, this system is aimed at the sent warning which is estimated from the monitoring results.

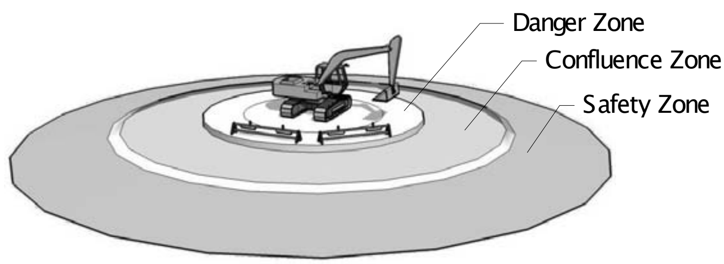


Figure-1. Defined area image for warning

Data Input

On construction sites, workers of different affiliations work at their individual tasks. Thus, to enable individual workers to recognize the danger to their particular situation, it is necessary to optimize the warning contents for each worker. This system defines persons, heavy equipment, and administration as three categories that receive a warning [Figure-3]. Persons are divided into two, relevant and nearby, and spotters and signalmen were also added.

Individual information about participants is inputted to the warning system in order to select a suitable warning for the worker judged to be in danger. The individual information is classified into general and concerned. The general information specifies the individual who receives a warning, and the concerned information is limited to participants who supervise individuals in danger.

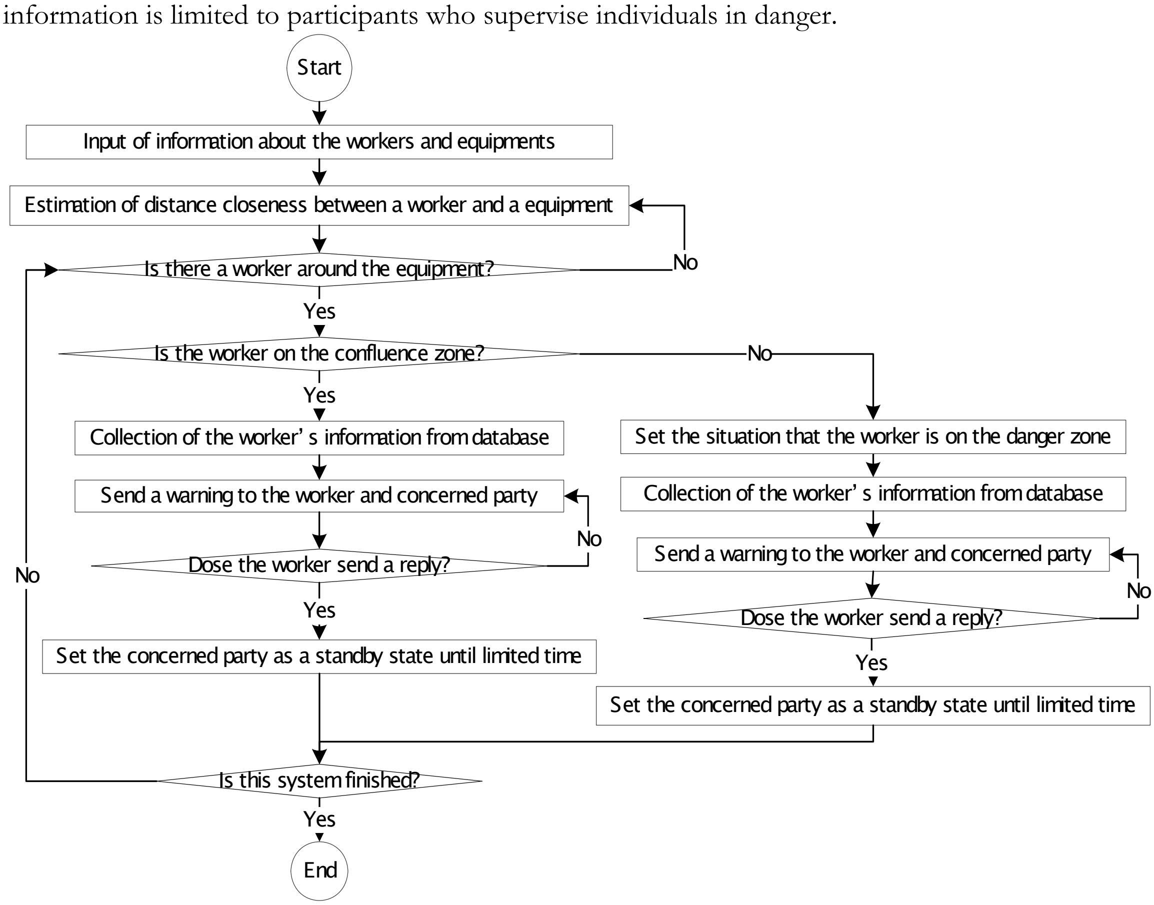


Figure-2. System flow to send warning

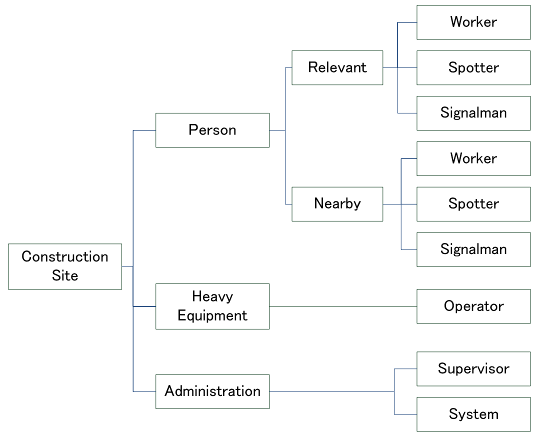


Figure-3. Classification of system participant

Table-1, Participants and input data

Participant	Input data	
Person	General	Name, Tag ID, Device address, Affiliation,
	Concern	Equipment model number
Heavy Equipment	General	Model number, Tag ID
	Concern	Operator's name Spotter' name Signalman's name
Administration	General	Device address
	Concern	Database code

Estimation of Distance from Equipment

The distance from the worker to the heavy equipment, which is used to judge the collision potential, is estimated from the working area information. The active type RFID was selected as the method of working area collection.

An RFID device consists of readers and tags. It periodically emits an identification number using a battery, so the object to which the tag is attached becomes identifiable. The tag is given to the worker and to the heavy equipment that needs working area information. The electric wave of a tag may be reflected and absorbed by the body of the equipment depending on the frequency band. In order to reduce the number of warning malfunctions, it is also necessary to install two or more tags.

The RFID reader receives the ID number from an RFID tag that exists within the range of the recognition area. The size of the recognition area varies according to the radio signal strength. The existence area of the object that has the received ID number is determined from the installation position of the RFID reader. The installation positions are determined according to the plane of the actual site. If the number of installation readers is large, it is assumed that the estimation accuracy is improved, but this would involve an increase in cost and in volume of data to be processed.

The working area of workers and equipment is estimated to exist in the vicinity of the reader's position using the ID data collected from each RFID reader. However, the estimation result has noise due to the electric wave characteristics. The author proposed a noise reduction method using RSSI (Radio Signal Strength Indication) [3][4].

The distance from equipment to a worker is estimated from the working area data. When the estimation result indicates that the worker is in the same working area as the equipment, the worker is treated as being in a danger area. When the estimation result indicates that the worker is in an area adjoining the equipment working area, the worker is treated as being in the confluence zone. It is necessary to set the boundary of the adjoining working area according to the moving speed of the heavy equipment.

Definition of Warning Message

The role of a warning is to greatly decrease the possibility of accident to a worker and related parities. Generally, warning lights and beeping sounds are used on construction sites. A worker who perceives a warning light or beeping sound is implicitly forced to get away from the danger, but there is insufficient information about what is going on around him. It is possible to reduce a worker's inadvertent error when trying to avoid the danger by providing him with information about the situation. In this system, a warning message is defined as a means for sending information about danger to a worker and related parties. The warning message is sent from an administration server when a worker enters a confluence zone.

The warning message to the worker was set up for four cases.

Case-1). A worker who is related directly to the heavy equipment work is in the confluence zone.

Case-2). A worker who is not related to the heavy equipment work is in the confluence zone.

Case-3). A response to a warning is sent.

Case-4). A response to a warning is not sent.

Participants who receive a warning are set up for each case, and the contents of the warning message are prepared accordingly. It is then checked whether the participants need to send a response when they receive a warning message.

Figure-4 and Figure-5 show examples of warning messages concerned with hydraulic excavation for case-1 and case-2.

The relevant worker, operator, and system are set up as the participants in Figure-4. The warning message of ["(worker's name), you are near the (equipment' name). Be careful to avoid being struck."] was prepared for the relevant worker. The worker's name and the equipment's name in the message are extracted from the database that inputted the individual information. The worker called by name can improve his consciousness of warning and confirm the situation from the equipment' name. The relevant worker who received the warning message needs to reply to the system administrator.

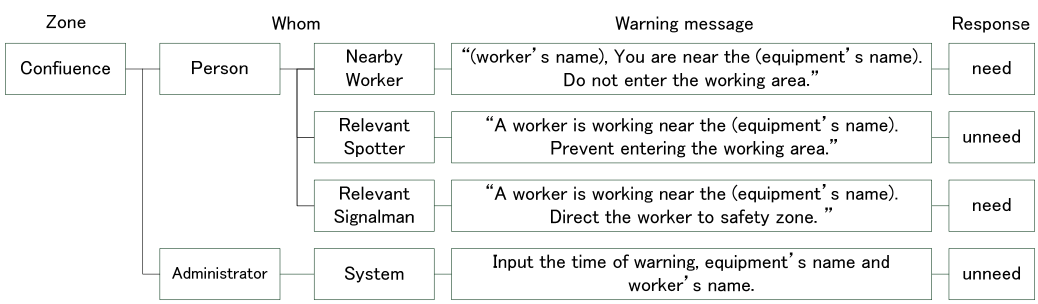


Figure-4. Example of warning message on the case-1)

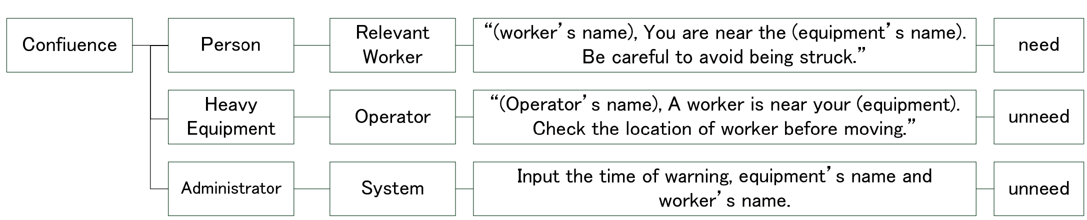


Figure-5. Example of warning message on the case-2)

Development of Warning Device

Basic Concept

The warning system that sends the message is composed of four parts: an active type RFID reader, a worker, a piece of heavy equipment, and an administrator server [Figure-6]. Since the worker and equipment are moving on the site, each part transmits and receives data through a wireless network. A smart-phone was selected as a personal device for receiving the warning message from the administrator server, and for sending the response. A smart-phone is a device for handling several media formats such as voice and text via the telecommunications network.

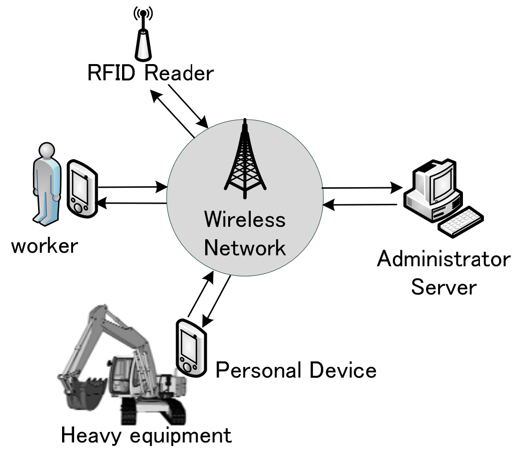


Figure-6. Collision-warning system component

Prototype

An experimental prototype for sending a message from the server computer to the smart-phone was made under the wireless local area network. The operating system of the smart-phone was Window Mobile 6 Professional, and Microsoft Visual Studio 2005 was used as a software development environment.

The administrator server detects the worker who is in a dangerous situation second by second, and sends a signal that includes a command to play the warning message. Usually, voice and letter are used as the information media, but this system adopted voice media so that the gaze duration for confirming the message is less than that of the letter media. The network load is reduced by preliminarily downloading the audio files of warning message and sending only the IDs of file to each device.

A wireless headset for hearing a message and replying to it is used by Bluetooth network. There is a possibility that manual operation of the device could create an accident situation. Therefore, it is not to stop working hands necessary for operation. The voice-recognition function that is fitted into a smart-phone helps device operation such as incoming signals and calling from spoken worlds.

Conclusion

The basic concept and functions of a warning system for preventing collision accidents was proposed. The results show that the warning system optimized for individual people is efficient in reducing a worker's inadvertent error, recognizing warning information sharing by concerned parities, and confirming avoidance. A smart-phone was selected as a communication device for receiving messages and responding to them. Acceptance of a voice message and responding via a wireless network was verified by a prototype.

Acknowledgement

This work has been supported by a Grant-in-Aid for Scientific Research on Priority Areas (B) No.20360282 of the Japan Society for the Promotion of Science (JSPS).

References

[1] Japan Advanced Information Center of Safety and Health , (Webpage 2008.10.23) http://www.jaish.gr.jp/information/sokuhou.html
[2] Soungho CHAE & Tomohiro YOSHIDA (2008), A study of safety management using working area information on construction site, 25th ISARC, 292-299.
[3] Soungho CHAE & Naruo KANO (2005) A location system with RFID technology in building construction site, 22nd ISARC, 43(1)-(6).
[4] Soungho CHAE & Naruo KANO (2007), An estimation method of working area with RFID technology, 5th ICCPM/2nd ICCEM, 47(1)-(8).

Detecting and locating leaks in Underground Water Mains Using Thermography

Mohamed Fahmy, P. Eng., M. ASCE[1] and Osama Moselhi, P. Eng, F. ASCE[2]

1PhD candidate, Dept. of Building, Civil, and Environmental Engineering, Concordia Univ., 1455 Blvd. de Maisonneuve W, Montreal, QC, H3G1M8.Tel: (514)848-2424/3902 Fax: (514) 848 7965, E-mail: ma_fahmy@encs.concordia.ca
2Professor, Dept. of Building, Civil, and Environmental Engineering, Concordia Univ., 1455 Blvd. de Maisonneuve W, Montreal, QC, H3G1M8.Tel: (514)848-2424/3190 Fax: (514) 848-7969, E-mail: moselhi@encs.concordia.ca

Abstract

Water systems all over the world experience water losses. Leakage is the most common reason of water loss. Problems associated with water main leaks are a growing concern around the globe. These problems include water and energy loss, in addition to considerable properties damage. In current practice, not all water leaks can be detected due to intensive time and expensive cost associated with the leak detection process; consequently, some leaks are still occurring and lead to problems mentioned above. Management of water leaks can be improved if leaks can be detected effectively then rectified efficiently. This paper presents a study conducted for detection of water leaks, and identification of their respective locations in underground pipelines using Thermography IR camera. The paper describes the field work and the experimental protocol, which were carried out over two years in three different locations in greater Montréal (Canada) area in order to investigate factors that affect the applicability and limitations of the IR technology used in this study.

Keywords: Water pipelines; Thermography; Water Leak detection; Failure investigations

Introduction

Thermography (IR) camera measures and images the emitted infrared radiation from an object. It can detect thermal contrasts on pavement surface due to water leaks. In addition, it enables relatively large areas to be investigated effectively in less time and consequently less cost comparing to currently leak detection methods. It is also independent of pipe type and size. Also, it can be used in day or night time. These advantages make using IR camera overcomes limitations associated with currently leak detection methods. This paper presents a study conducted to investigate the factors that affect the use of IR camera in detecting and locating water leaks in underground water mains such as weather conditions, soil and pavement surface conditions, ground water level and distance of sensor (i.e. IR camera) from source. The study also focuses on the impact of camera setting and vehicle speed on which the camera is mounted on the accuracy of the results obtained. Case example is presented to illustrate the use of the proposed methodology.

Proposed Methodology

The methodology presented in this study is based on intensive literature review, meeting with experts, and on the analysis of actual data collected from three municipalities in greater Montreal area; Pierrefonds, southwest, and downtown Montreal (Canada). The development of the methodology involves five major steps: I) identification of factors that affect thermal contrast at pavement surface; II) field investigation and on site experimental work; III) analysis of the data obtained in order to determine the most suitable conditions of using IR camera for the detecting and locating water leaks, IV) establish the relationship between the detected leakage area at pavement surface and the location of leak in the water main being tested; V) validation of the proposed methodology by comparing leak locations detected by the proposed system and by acoustic-based methods. Figure 1 shows the proposed methodology.

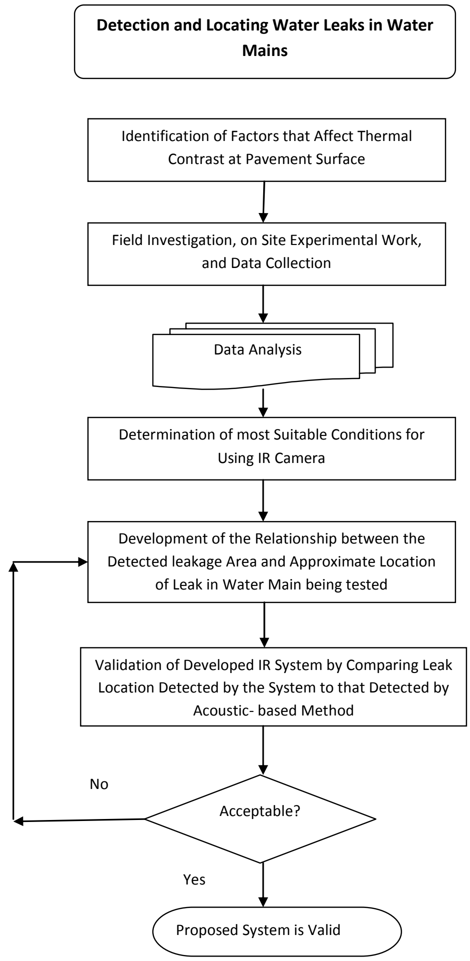


Figure 1: Proposed Methodology

Thermography (IR) Camera System

The ThermaCAM S 60 infrared condition monitoring system was used in conducting a set of field experiments. The system consists of an infrared camera with a built in 24o lens, a visual color camera, a laser pointer, and infrared communications link (FLIR SYSTEMS 2004). This system provides real time high resolution color images in both infrared and visual modes. The visual mode was used to check the existence of any foreign bodies on the pavement surface, which might affect thermal contrast. To document the thermal variation on pavement surface due to water leaks of pipes below ground it is possible to capture and store images on a removable flash card. The captured images that have sequential numbers can then be analyzed in the field using the developed methodology to determine approximate locations of leaks. Figure 2 shows the thermal contrast between areas with lower temperature (dark areas) that represent pavement surface temperature in its natural state (i.e. without leaks) and the bright areas that indicate water leak.

I. Factors That Affect Thermal Contrast at Pavement Surface

Based on a comprehensive literature review, seven interviews with experts, and preliminary field investigation it became clear that the following two major factors affect the thermal contrast at pavement surface.

1. Heat Balance at Pavement Surface

At the top of the pavement surface, four modes of heat transfer are considered: conduction into the pavement layer, convection, solar absorption, and grey-body irradiation to the surrounding (ASHRAE 1981; Hutcheon and Handegord 1983; Bentz 2000, Schlangen 2000). For irradiative heat transfer at the top pavement surface, two contributions are considered the first is radiation absorbed from the incoming sunlight. The second is the emission of radiation from the pavement to the sky (McCullough et al. 1999, Loomans et al. 2003).

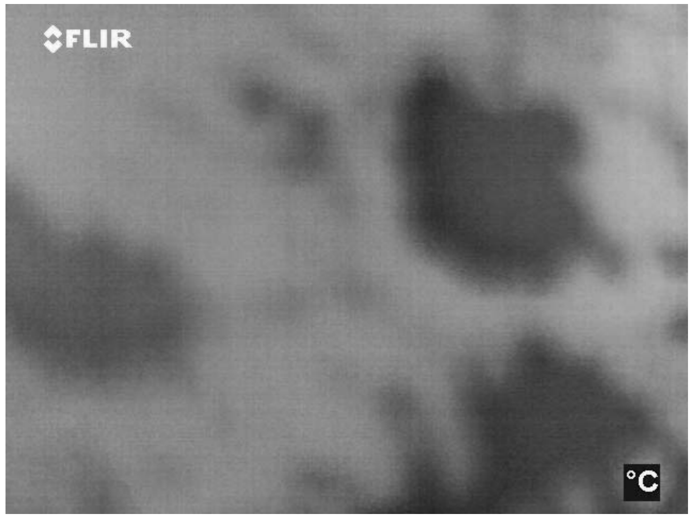


Figure 2: Thermal Contrast at Pavement Surface due to Water Leak

2. Heat and Moisture Transfer in Soil

Migration of heat and moisture in soil is a coupled energy and mass transport process, which is affected by the field distribution of temperature, pressure, and velocity (Liu et al. 2005). Soil heat transfer in the unsaturated zone of the soil is the sum of fluxes due to heat conduction and convection (Sung et al. 2002).

II. Field Investigation and on Site Experimental Work

Field investigation was conducted using thermography IR camera and the results obtained were compared to the results obtained using acoustic-based system, which will be referred to in this paper as "leak finder". Leak finder locates water leaks by detecting the sound or vibration induced by water leaking from pressurized pipes (Hunaidi et al. 2004).

Preliminary field experiments showed that the moisture level near the pavement surface affects the pavement surface temperature because of its major influence on the thermal properties of the soil. Furthermore, it was found that the thermal contrast detected by IR camera was close to the exact location of the leak detected by the acoustic-base leak finder device. Following the preliminary survey, detailed field investigation and experimental work was conducted in order to determine the thermal performance of water leaks in underground pipelines, and establish relationship between the detected leaking areas and the accurate location of the leaks.

In order to attain these objectives 42 water pipelines were scanned using IR camera. The diameter of these pipes ranged from 150 to 200 mm and their length ranged from 48m to 300 m. The field tests were conducted in down-town Montreal, South-West Montreal, and the Pierrefonds municipalities in Canada. The study presented in this research was carried out over 24-month period from July 2005 to August 2007, and the timing of the fieldwork was selected to represent a wide range of weather conditions in terms of prevailing light, and ambient air temperature. The inspection was also executed throughout a range of cloud cover from clear sky to overcast. This allowed testing the effectiveness of energy transfer between the sky and the investigated pavements. It should be noted that, the measured temperatures using IR camera were compared to those measured using thermocouple device. The average difference in measured temperature was (+/- 2 ^{0}C).

In order to obtain obvious color contrast in acquired images, IR camera set up was adjusted based on number of trials. The distance from the pavement surface to the camera ranged from 1.20 m to 12.0 m. Combinations of various ranges of vehicle speed, on which the camera was mounted, and time intervals of image capturing were carried out. The vehicle speed ranged from 5 km/hr to 20 km/hr and the rate of capturing images ranged from image/2 sec to image/10 sec.

III. Analysis of Data Obtained

Figures 3 shows the relationship found between pipe temperature, average ambient air temperature and average pavement temperature in fall season. As shown in Figures 2 the warmer pipe temperature indicates high possibility in detecting water leaks using IR camera according to Equation 1 (ASHRAE 1981; Hutcheon and Handegord 1983; Bentz 2000)

$$Q_{cond} = K_{cond} * (T_p - T_s) / L \qquad (W/m2) \qquad (1)$$

Where K cond is the average thermal conductivity of the soil and pavement in (w/m. K), Tp and Ts are the pipe temperature and surface temperature respectively and L is the length of the flow path (i.e. burial depth).

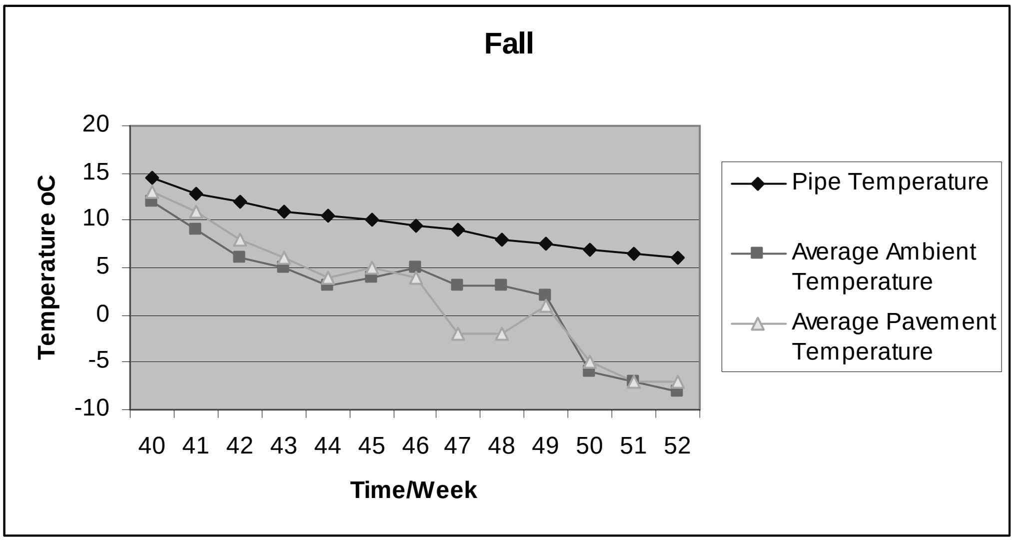


Figure 3: Comparison between Temperatures in Fall Season

Cloud Cover and Prevailing Light

Data collected in this research showed that pavement temperatures under clear sky and/or during day time were consistently warmer than pavement under cloudy condition and/or at night and early morning. As a result, detection of leaks will be more accurate under overcast condition between 11 pm and 6 am.

Change in Thermal Characteristics of Soil and Pavement Surface

Soils close to water leaks experience increase in moisture content and may become saturated. Such change in moisture content changes the thermal characteristics of the soil and makes it more conductive to heat relative to dry soil away from the leak. The soil temperature variation observed in this research through the four seasons indicates that the soil temperature presents higher variation in shallow than in deeper depth. During winter, the average soil temperature of deeper layers is higher than that at the shallow soil depths. It means that, during winter the heat is transferred from the deeper soil depths to surface, while during the summer months it changes direction. This is in agreement with the recent study conducted by Antonopoulos (2006). Also during the winter period, the rate of change in soil temperature under snow cover was less due to low thermal diffusivity and high albedo of the snow. It was found that areas detected that have a thermal contrast on pavement surface decreased with soil surface evaporation during daytime and slightly increases during nighttime.

Infiltration into Adjacent Sewer Pipes

The field investigation and experimental work carried out in this research revealed that more than 40% of the water leaks detected were infiltrated into adjacent sewer pipes, that prevent the moisture movement from reaching pavement surface, consequently, that type of leaks could not be detected using IR camera.

Ground Water Table

The ground water table has a great influence on the use of IR camera; experimental work conducted in the vicinity of Saint Laurence River in Montreal showed that the ground water table was higher than the pipe level. The IR images captured to the pavement surface at this area showed no variation in the thermal properties of the pavement surface.

Distance of Sensor from Source

The impact of the distance from the pavement surface to the camera was studied. Tests were conducted over a range from 1.20 m to 12.0 m. The experimental works revealed that the more the distance between sensor (IR camera) and pavement increases the more the thermal contrast enhances and vice versa. Therefore, more distinction of leakage area was obtained from a distance of 12.0 m from the pavement surface.

Vehicle Speed and Rate of Capturing Images

Combinations of twelve sets represent various ranges of vehicle speed and periodic capturing of images was carried out. The vehicle speed ranged from 5 km/hr to 20 km/hr and the rate of capturing images ranged from image/2 sec to image/10 sec. the best results obtained in terms of distinguishing thermal contrast and accuracy when the vehicle speed was set at 5 km/hr and the rate of capturing images was set at image/2 sec.

Effect of IR Camera Setup

In order to obtain obvious color contrast in acquired images, camera set up was adjusted based on sets of thirty six trials, it was found that the emissivity, palette type, and noise reduction function that reduce clutters, were the most effective parameters. The final selection of these parameters was as following: emissivity was selected based on pavement status ranged from 0.85 for snow cover, 0.90 for dry pavement surface and 0.94 for wet pavement surface. Palette iron was selected, which provides finest contrast and color degradation ranged from blue (i.e. represents lowest temperature) to white (i.e. represents highest temperature). Noise reduction function was activated.

IV. Modeling

The approximate locations of water leaks carried out in this research are based on two major steps:

1. Determination of areas that indicating thermal change at pavement surface (i.e. water leaks)
2. Establish the relationship between detected leaking area and pipe burial depth

1- Determination of Areas Indicating Water Leaks

Twenty five pipelines experienced water leaks were tested using IR camera in order to develop Equation number 2. Applying Equation number 2 the approximate location of leak can be found. Then, the user has to move to that location to determine the entire area that experience thermal contrast.

$$X = \frac{(N-1) \times 0.28\, S}{R} \qquad (2)$$

Where:

X: approximate location of water leak from the origin point (m)

N: chronological image number

S: average vehicle speed (Km/h)

R: rate of capturing IR image (image/sec)

2- Establish the relationship between detected leaking area and pipe burial depth

Field observations conducted in this research revealed that the detected thermal contrast due to water leak on pavement surface approximately represents a circular base of a cone, which its head represents the location of the leak in the pipe being tested.

V. Validation of Proposed Methodology

The leak locations detected using IR camera for twenty five water leaks were compared to those detected using the acoustic-based leak finder method. The results are shown in Figure 4. As shown in Figure 4 the difference ranged from 1.01m to 2.30m

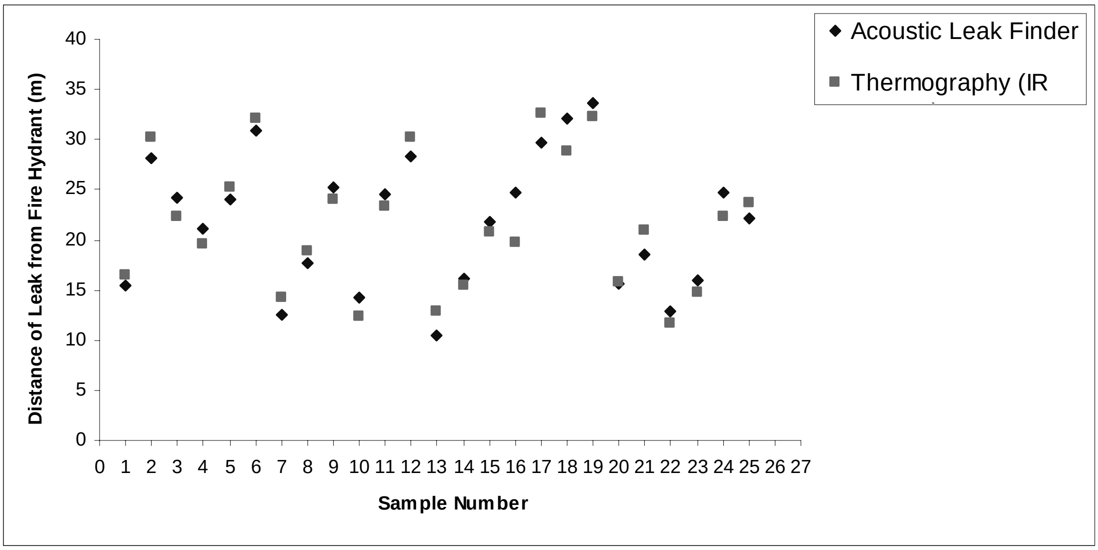


Figure 4: Leak Locations (m) using Acoustic Leak Finder Verses IR camera

Case Example

In this example we consider a 6" diameter CI water main segment, 48.7 m length (i.e. the distance between two fire hydrants), located 1.80 m below the ground in a residential area of the city. The municipality asks for performing a condition assessment work on that pipeline using thermography method (IR camera) and verifying the results by using the acoustic-based leak finder. The inspection team utilized a vehicle with speed 6km/hr and the rate of capturing images was 0.5 image/sec.

Applying the methodology described above the user found that

1- image number 6 showed thermal change as shown in Figure 5
2- The user moved to the location represents image number 6 and perform further investigation.

Appling Eq. 2

$$X = \frac{(N-1) \times 0.28\ S}{R} = \frac{(6-1) \times 0.28 \times 6}{0.50} = 16.80 \text{ m}$$

The results obtained is shown in Figure 5

Summary and Conclusion Remearks

This paper presented a study on the use of Thermography IR camera for detecting and locating leaks of water mains. The study encompassed field investigation and testing as well as modeling development. The field work was conducted on water mains in 3 locations in the greater Montreal area. The IR camera detected successfully number of leaks as a thermal contrast at pavement surface that occurred in fall and spring seasons, while it failed in detecting leaks occurred in summer and winter due to high pavement temperature and the snow coverage, respectively. The thermal contrasts due to water leaks take a shape of near circular cone base. The head of the cone represents the approximate location of leak. However, using IR camera in vicinity of sewer pipe was not reliable. The near optimum diurnal time of using the camera was between 6-8 am. The leaks detected using IR camera was compared to those detected using acoustic- based leak finder method. A case example is presented to demonstrate the use and accuracy of the developed methodology.

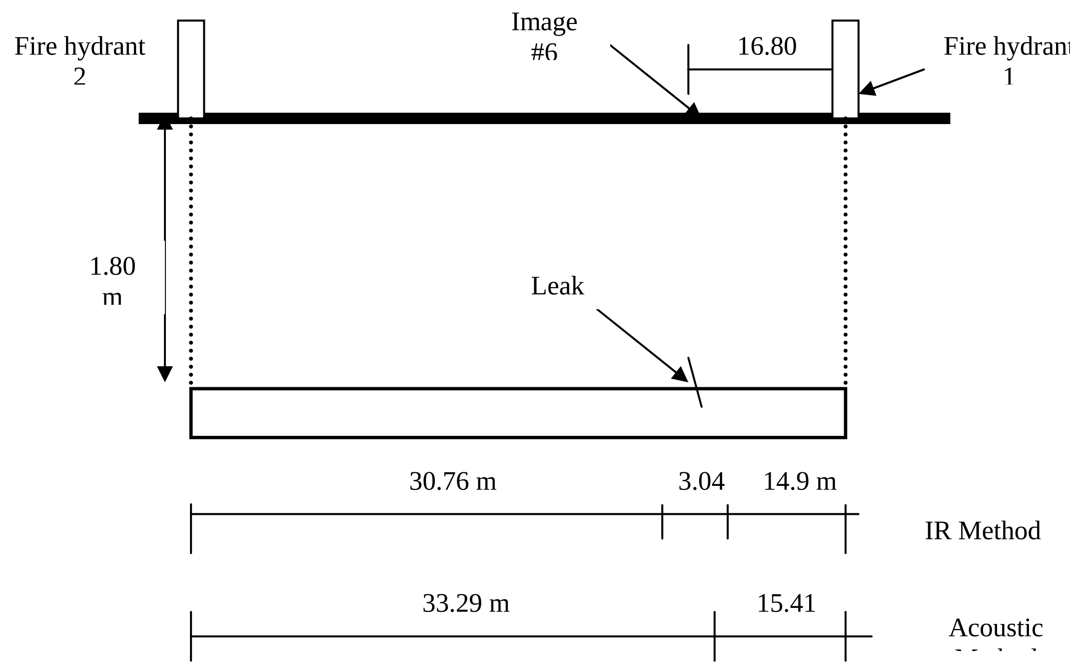


Figure 5: Case Example

Acknowledgement

The authors whishes to acknowledge the financial support provided by the Natural Sciences and Engineering Research Council of Canada, and the internal research grant provided by the Faculty of Engineering and Computer Science, Concordia University.

References

[1] Antonopoulos, V. (2006). "Water Movement and Heat Transfer Simulations in a Soil under Ryegrass" Journal of Biosystems Engineering, 95 (1), pp 127–138.

[2] ASHRAE, (1981), ASHRAE Handbook Fundamentals, NY.

[3] Bentz, D., (2000), " Acomputer Model to Predict the Surface Temperature and Time-of- Wetness of Concrete Pavements and Bridge Decks", National Institute of Standards and Technology, US Department of Commmerce.

[4] FLIR Systems AB. (2004). ThermaCAM S60 Operator's Manual.

[5] Hunaidi, O.; Wang, A.; Bracken, M.; Gambino, T.; Fricke, C. (2004)." Acoustic Methods for Locating Leaks in Municipal Water Pipe Networks" International Water Demand Management Conference, Dead Sea – Jordan.

[6] Hutcheon, N., and Handegord, G. (1983). Building Science for a Cold Climate. John Wiley& Sons, Toronto.

[7] Liu, B, Liu, W., Peng, S., (2005). "Study of Heat and Moisture Transfer in Soil with a Dry Surface Layer", International Journal of Heat and Mass Transfer Vol 48, pp. 4579- 4589.

[8] Loomans, M, Oversloot, H, De Bondt, A, Jansen, R and Van Rij, H, (2003). "Design Tool for the Thermal Energy Potential of Asphalt Pavements", Eighth International IBPSA Conference, Eindhoven, Netherlands.

[9] Mcullough, B., and Rasmussen, R. (1999). Fast Track paving: Concrete Temperature Control and Traffic Opening Criteria for Bonded Concrete Overlays. Federal Highway Administration, Mclean, VA.

[10] Schlangen, E, (2000). Heat of FEMMASSE Manual. INTRON BV, The Institute For Quality Assessment In The Building Industry In The Netherlands.

[11] Sung, K., Yavuz, C., Drew, M. (2002). "Heat and Mass Transfer in the Vadose Zone", Journal of Contaminant Hydrology, Vol. 57, pp 99-127.

Development of Digital Photo System Using RFID Technology In Plant Construction Management

Hatori Fumio[1], Yoshimura Yasushi[2] and Ebata Shinichi[3]

1:Hitachi Plant Technologies,Ltd. Business Development Div. 5-2 Higashi-Ikebukuro 4-chome,Toshima-ku,Tokyo,170-8466 Japan, e-mail:fumio.hatori.zq@hitachi-pt.com
2:Hitachi Plant Technologies,Ltd. Matsudo Research Laboratory
3:Hitachi Plant Technologies,Ltd. Energy Systems Div.

Abstract

We have studied the applicability of RFID tags and developed systems to provide clear work traceability to streamline administrative task in the construction work of a plant. Our test results indicate that RFID tags can be used effectively for these purposes. We have developed a method for effectively proving whether a digital photo is authentic or not in order to obtain clear work traceability, along with a system to reduce the time spent filing digital photographs and associated information correctly in a short time.

Keywords: Material management, progress management, IC tag, digital camera

1. Introduction

The rationalization of construction costs is strongly urged. On the other hand, quality control and the safety management of the construction become very important, too. But, the labor productivity of construction industry is low in comparison with the other manufacturing industry. In order to solve these problems, construction management system using information technology becomes more important. RFID (Radio Frequency Identification) is attracting increasing attention as a basic technology for the incoming ubiquitous society. Various experiments are being performed to apply this technology to commodity distribution businesses and production plants. For the practical use of an information system, there is also the need to develop underlying hardware as well as useful applications.

The construction of a large-scale plant requires a great number of equipment, materials, workers and temporarily stored products. In order to assure the construction quality, the traceability of work history, from factory production up to an installation inspection in a destination site is required so that people can find who did particular works when and how, which involves enormous amounts of manpower management. In this report, we study the environmental resistance of RFID in a plant construction site, and work traceability as the key control item to view the applicability of RFID, and describe the development of RFID-used applications.

2. Plant construction Management and study of RFID applications

2.1 Traceability of work

To assure the quality of construction work, a work history is kept to see when and who did which task, as well as recording working conditions and inspection results. This operation is called work traceability, providing a work quality assurance record and the means of finding the root cause if a defect is found after a plant starts operation.

Fig. 1 shows the flow of example installation work of pipe grooves matching and work records. Pipes delivered to a building interior are temporarily suspended near their installation site. An inspection is performed to check that no foreign matter is contained in the pipe interiors, and photos are taken to prove it. Subsequently matching of pipe grooves is performed. During this stage, when the work was carried out and by whom is recorded, according to the instructions for grooves matching work. Upon completion of this work, a witness inspection is conducted to see and record whether the work has been done as instructed. If the work passes the inspection, welding work is performed. In addition to the next welding and piping works, other works are also recorded to ensure work quality and provide traceability for the same. These control subjects can be identified by RFID and their information can be input by RFID, helping to improve recording efficiency and quality control.

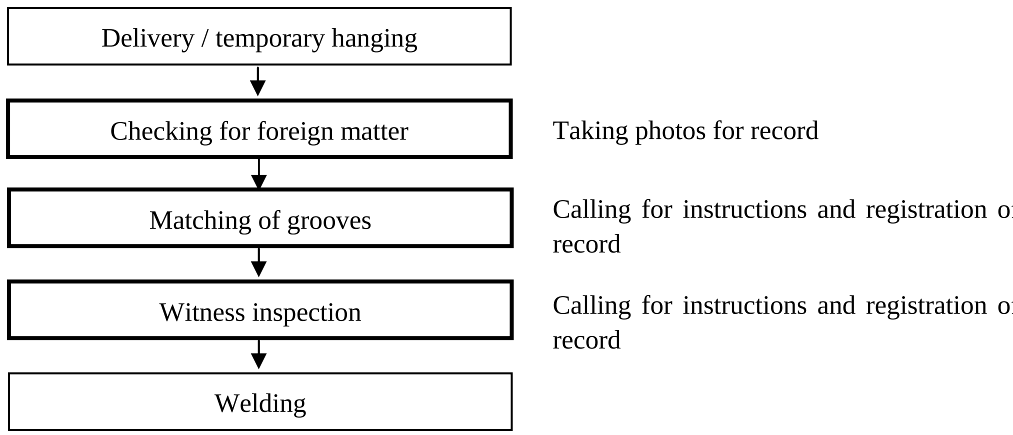


Fig. 1 Piping work and recording

3. Problems for RFID applications and study method

3.1 Assessment of environmental resistance

Assuming the use of RFID in construction sites, we have selected the assessment items of RFID environmental resistance, which are indicated in **Table 1**, along with the assumed conditions. We have conducted tests to check that RFIDs can perform reading without being damaged and to see whether an RFID can be used for construction work without any problem. In addition, low cost is a prerequisite because a huge quantity of RFIDs may be used for material control. Out of the RFIDs available when we performed assessment, we selected widely-used low-cost RFIDs and tested them. **Fig. 2** shows the conditions of the test conducted to assess the influence of welding heat on RFIDs. The RFID tags were exposed to the simulated environmental conditions of construction sites in order to test them for the respective assessment items and check whether an RFID reader can read the IDs recorded in the tags.

The attaching of RFID tags at equal intervals from the weld line: Measurement of temperature with a thermocouple

Table 1 Assessment items of RFID environmental resistance and assumed conditions

No.	Environment of the construction site		Assumed conditions
1	Temperature	At high temperature	Welding is performed near RFIDs
		At low temperature	Stored outdoor in winter (-20°C)
2	Noise influence		Use of RFIDs near a welding machine
3	Transportation of products		Shocks are given when products are transported and unloaded.
4	Hydric environment	Rain	When it is raining
		Snow	When it is snowing
5	Contamination	Iron powder	Iron powders are adhered to an RFID in a building under construction.
		Paint	Paint is adhered to an RFID.

Fig. 2 Assessment test of welding heat

3.2 Problem in work traceability

Documentary photography is vital as evidence for construction work traceability. However, the present documentary photography has the problems shown in **Table 2**.

There are many similar-looking inspection subjects such as pipes, and so the picture of a subject is taken together with a board on which the necessary information is written so that the subject can be identified. If no such board is photographed, a documentary photo cannot specify which product is shown. However, this method cannot prove the information given by a photo is the same as that given by a board. Furthermore, when organizing many photographs, photo files must be opened one by one to check the board information, which is a time-consuming operation.

To solve the above problems, we have studied methods to prove whether a photo is authentic or not and reduce the time required to organize photos and search them, and developed an effective RFID-applied photography control system.

Table 2 Present problems for photography control

Item	Present status	Problems
To prove whether a photo is authentic or not	Information of a subject is indicated on a board and photographed.	It is difficult to prove whether the board information is really for a taken photo.
Efficiency for organizing photos	Screen images should be checked to see photo images.	A person should see a screen image and judge.
Efficiency for searching a photo	Photos are placed in work record control files.	It takes time to find a specified photo.

4. Results of RFID applicability assessment and development of a new application

4.1 Results of the environmental resistance assessment

Table 3 shows the environmental resistance test results of RFID. ○ indicates that the RFID tag was not destroyed and could be read quite normally. △ indicates that the reading distance was shortened and × indicates that no reading could be made.

In most test items, not all the tags were destroyed and some could still be read. However, if they were remained wet or in direct contact with metal, reading could not be made, although when tags are used outdoors in rain or snow, the results show that they can be read without any problem if moisture is removed. Likewise, if an isolation material is used to separate a tag from a metal by 2~5mm, readings can be made. In a test for the transportation of products, 2.45GHz and 950MHz tags with a dipole antenna were found to be readable provided their chips were not destroyed even if their antennas were cut. However, the 13.56MHz coil-shaped tag could not be read if its antenna was cut. The antenna of the 2.45GHz tag is smaller than that of the 13.56MHz tag, and so the 2.45GHz tag is less vulnerable to an impact. In particular, when a tag is attached to a heavy material such as a pipe, a bigger tag is more likely to be subjected to impact and damaged due to its size.

Table 3 Assessment results of RFID environmental resistance

No.	Test environment		Type（Frequency）13.6 MHz	2.45 GHz	950 MHz	Remarks
1	Temperature	At high temperature	○	○	○	Temperature near RFID: 283°C
		At low temperature	○	○	○	Reading can be made at - 20°C
2	Noise influence		○	○	○	Welding machine
3	Transportation of products		×	△	△	Capability is reduced when the antenna is cut.
4	Hydric environment	Rain	○	○	○	Usable with an additional measure
		Snow	○	○	○	Same as above
5	Contamination	Iron powder	○	○	○	Usable with an additional measure
		Paint	○	○	○	

The RFID communication distance is also influenced by the output of an RFID reader. In a simulation for practical use, we used a low-output handy reader. The results indicated that the reading distance was shorter than that shown in the catalog, even under normal conditions. It was 50mm for the 13.6MHz tag, 30mm for the 2.45MHz tag and 100mm for the 950MHz tag. The difference between the actual values and catalog values is excessive, meaning the communication distance should be slightly longer for practical use.

The above results indicate that RFID can be used for a plant construction if an improved operation method is devised for a problem such as that caused by metal contact. And, we have found that an RFID may be damaged during product transportation and that a tag with a dipole antenna has higher damage resistance.

4.2 Development of a photography control system

Fig. 3 shows the outline of the RFID-applied photography control system. The RFID of a subject product such as a pipe is read by an operator in a construction site and the RFID information is sent to a digital camera. Subsequently the information transferred by the RFID reader is embedded into the photo file taken by the digital camera and the photo can be proven to be authentic. After returning to the office, an operator registers the photo files taken together with the RFID files into the database and can organize the photos and search for them by using the RFID codes. The following describes the detailed respective functions:

(1) Proof of photo authenticity

Fig. 4 shows the procedures used to record the photos. An operator for inspection and recording uses a computer system equipped with an RFID reader to read the RFID tag attached on their helmet and thus identify themselves. Subsequently the test results are recorded into the computer system, and the RFID tags used to identify the inspection subjects are read. The time (t_0) when the RFID is read is recorded in the computer system, whereupon the subsequent elapsed time is monitored. The time (t_1) when a photo is taken is checked, while the computer system also checks the time having elapsed since the RFID tag was read until a photo was taken (t_1-t_0) and judges whether the photo was taken within the allowable time (t_a) for taking photos. If a photo is not taken within (t_a), the read RFID tag information used to identify a product is deleted and the photo taken becomes invalid. If a photo is taken within (t_a), the RFID tag information is embedded into the photo information and recorded. The allowable time (t_a) should be the time during which it is impossible for an operator to move elsewhere to take a photo of other subjects. If a piece of RFID information is one and only and it is impossible to take a photo of a product other than the particular product in question, the RFID information of which has been read, we can prevent a mismatch between the RFID information and the subject in the recorded photo and prove the authenticity of the photo taken.

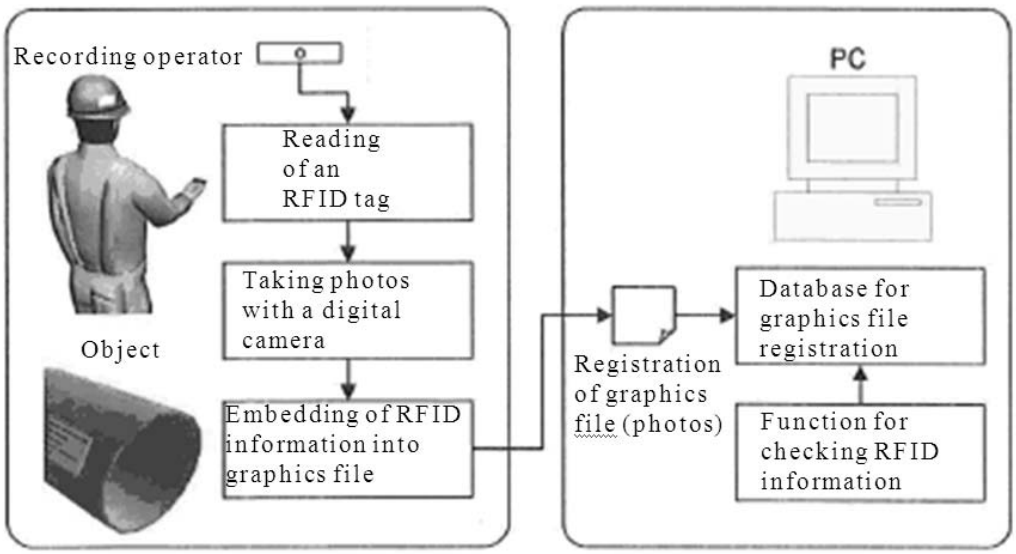
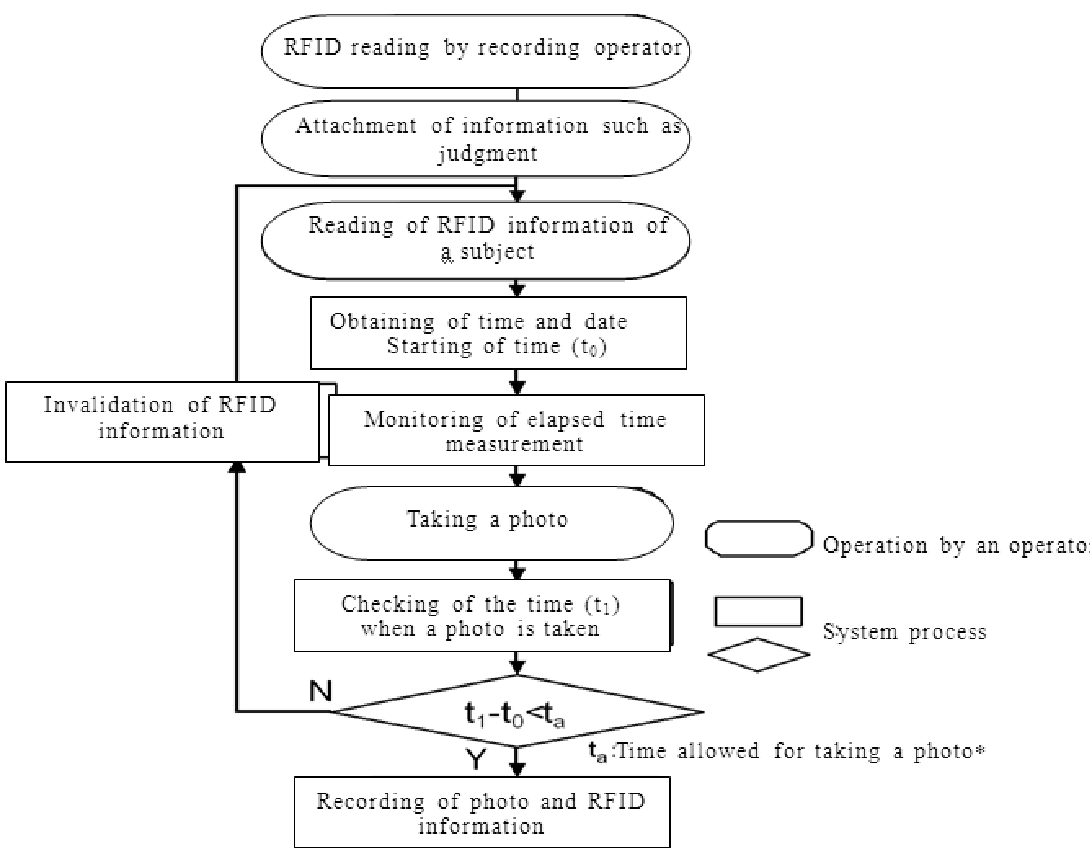


Fig. 3 Outline of an RFID-applied photography control system



*Time allowed for taking a photo: Allowable time for taking a photo after reading an RFID tag to prevent an operator from moving elsewhere or taking a photo of another subject.

Fig. 4 Procedures for RFID-used photo recording

(2) Filing of photos and searching function

We have studied methods for embedding an RFID code into a photo file to prove a taken photo is authentic and automatically judge a subject in a photo. The format of photos taken by a digital camera is standardized. The standard, known as EXIF (EXchangeable Image File format) specifies the types of information registered in a photo file and the storage location. As for the types of information, these include image size, resolution, photographing condition, maker name, camera model number, latitude, longitude and optional conditions, etc., while storage areas for writable information are specified for the respective types. A storage area for user information is also included in the same, and we have decided to use that area to write the read-in RFID code information.

Information written in EXIF can be referenced and checked in the application. Using the image database software, we have constructed a system where by referencing an RFID code recorded in the user information area of EXIF, an operator can check information associated with the RFID codes for subject items such as a recording operator and piping. This system enables an operator to search for the photos registered in it without opening files if the recording operator name or product name is given.

Fig. 5 shows the configuration of the RFID-applied photography control system. This system consists of a PDA and digital camera, both of which have the Bluetooth function, a short-distance radio communication standard. The PDA performs the input of an inspection record and reading of an RFID tag, and the data are transferred to the digital camera by Bluetooth communication in real time. RFID information is always transferred if a photo is taken within the allowable time. However, if no photo is taken within the allowable time, empty RFID information is transferred and the RFID information that has been transferred up to that point is deleted. The digital camera has a function to write RFID information sent by the PDA into the EXIF user information area.

For software to register and control the photos embedded with RFID codes, we have used the software for image control. This software is used to check the RFID code embedded in the EXIF user information area for photo files, search for specific photos and create a file sheet based on the RFID codes and information associated with the subjects.

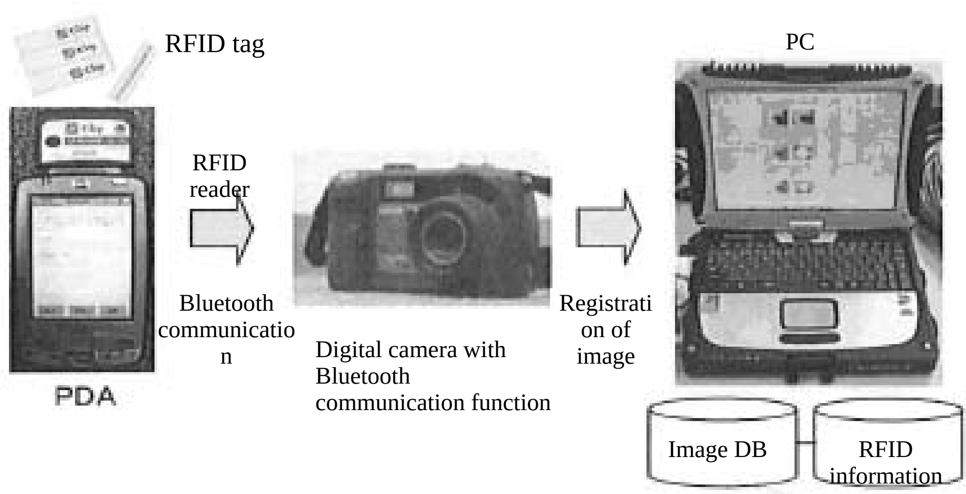
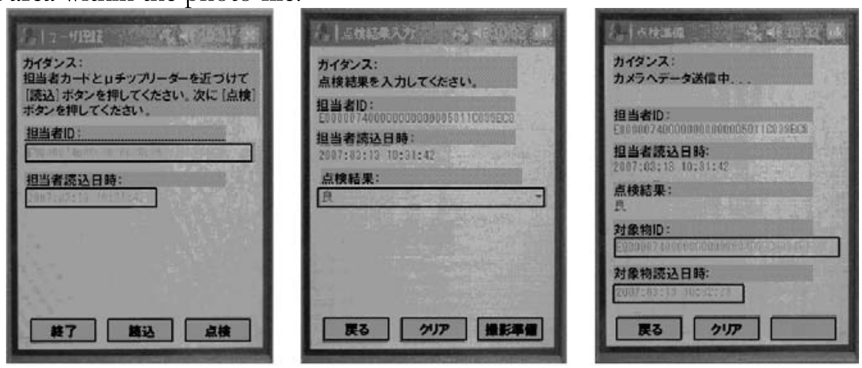


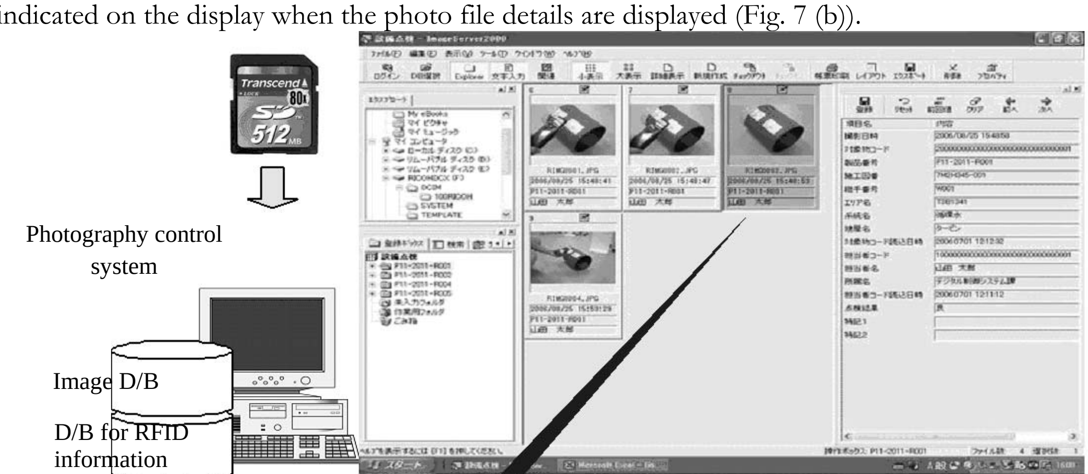
Fig. 5 Configuration of the photography control system

Fig. 6 shows the RFID-applied photography control system's screen transition. When the software is started, the user authentication screen is displayed to prompt an operator to read an RFID tag. The operator reads the RFID tag for self-recognition, which is attached to his helmet **(a)**. Subsequently a pass/fail judgment is performed in the inspection result input screen **(b)**. When the photographing-ready button is pressed, the RFID reader enters the state where reading can be performed. The reader reads the RFID tag attached on a subject pipe and transfers the RFID code and inspection result to the digital camera while keeping the read-in time of the tag as shown in the data transferring screen **(c)**. The transferred RFID code is displayed on the liquid crystal screen of the digital camera so that an operator can check the data transfer status. When a photo is taken in this state, the RFID code and inspection result are embedded into the EXIF user information area within the photo file.

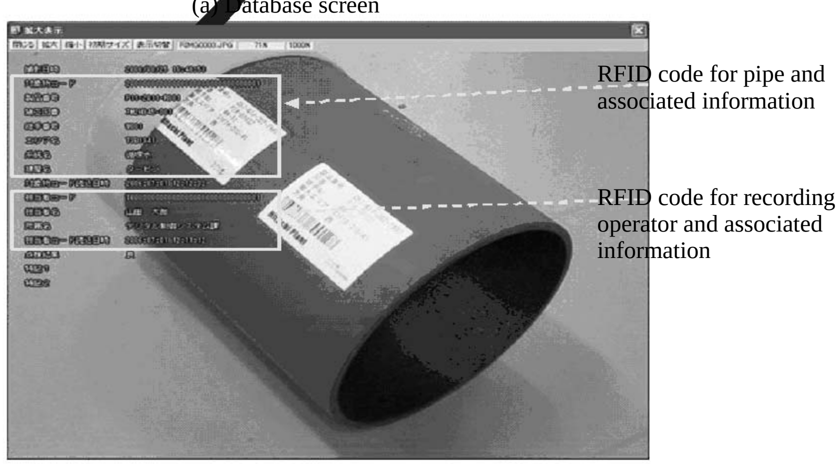


(a)User authentication screen (b)Inspection result input screen (c)Data transfer screen
Fig. 6 Diagram showing transition of PAD screens

A database software screen for photography control is shown in Fig. 7. The photos taken are recorded in a memory card, while the photo files recorded in the latter are registered in the image database. The belonging section, the operator's name for the operator RFID code and the drawing number, welding point number, system name and similar for the piping RFID code are recorded in an orderly manner within the database software in advance. The photo file is registered and at the same time the RFID information is called from the EXIF user information area to display the associated information. This information is indicated on the display when the photo file details are displayed (Fig. 7 (b)).



(a) Database screen



(b) Display of the photo file details

Fig. 7 Database software screen for photography control

c) Checking of the effects

A RFID code is issued on a one and only basis, meaning a subject pipe can be uniquely identified provided no human error is made when attaching the RFID tag to a pipe. Therefore, if the RFID code for the subject item is embedded into the photo without fail, the photo's authenticity can be proven. Only provided a photo is taken within the specified time, can this system allow registration of the RFID for the photo in order to definitely associate the photographic subject with the RFID code. We did set this allowable time for photographing at 30 seconds and tested photography accordingly. Under the circumstances simulating a construction site, an operator was well able to take a photo of a subject item within 30 seconds. However, it was impossible for an operator to move elsewhere to take photos of other subject items. For the construction of a plant in which bulky products are used and the photographic subjects are located over a wide area, the reliability of this system to prove whether or not a taken photo is authentic can be increased by restricting the time allowed for transferring RFID information. For small-diameter pipes, however, it was impossible to prove 100% correctly whether the photos for all the subject items were authentic or not. We will study the range of valid authenticity as well as the practical use of the system.

By using this system instead of the traditional method for writing RFID information on a board in a construction site, an operator need only read information from the RFID tags attached to the operator and product and take photos, meaning the work can be simplified and the working time reduced. Traditionally an operator had to spend as long as 2 to 3 hours checking photos and filing those taken for the purposes of product inspection work, which took half a day. We have confirmed that this system enables an operator to search for photos and file them automatically in an orderly manner within about 30 minutes.

5. Conclusions

We have studied the applicability of RFID technology to streamline the control work in plant construction, developed an RFID-applied application and obtained the following conclusions:

- We have conducted environmental resistance tests for RFID under the circumstances whereby a construction site is simulated It has since been confirmed that the RFID system could be used practically without a problem.
- We have proposed methods to embed RFID information into a photo file and restrict the time allowed to transfer RFID information
- By using a database system for photography control, we have developed a system enabling us to automatically call a subject item from RFID information and create a file sheet.

We have already started using our developed system, and confirming the application effect. We think the RFID technology progresses still more. We will also continuously check new products to improve it and study the application of this technology.

References

[1] Streamlining of on-site work by using short-distance radio wave technology written by Eiko Ashikawa and Yoshio Shida and published in 2004

[2] "Control system for on-site digital photography and design information by using IC tags" and "Technical papers for using civic work information" written by Rinsho Shiya and Nobuaki Yabuki, Vol. 14, pp. 59-66, published in 2005

[3] Digital still camera technology expert panel: JEIDA standard and Image file format standard for digital still camera (EXIF) Version 2.1 published by the Japan Electronic Industry Development Association in 1998

Human Motion Analysis Using 3D Range Imaging Technology

Ray Gonsalves[1] and Dr. Jochen Teizer[2]

[1]RAPIDS Laboratory, School of Civil and Environmental Engineering, Georgia Institute of Technology, Phone: +1-404-894-8269, Fax: +1-404-894-2278, e-mail: teizer@gatech.edu

Abstract

Human motion analysis and tracking are a significant research area in the domain of computer vision. Existing systems of today focus on detection of human targets by analyzing their movements in order to recognize the different activities performed by them. Our research work mainly focuses on using detection and tracking of human targets using a 3D range image camera [1] for surveillance purposes in order to ensure the safety of construction workers and also to monitor their posture and movements for heath related purposes in an active work zone. For this purpose the tracking algorithm proposed performs the segmentation of a human target (i.e. construction worker) from a range image video sequence and then models and tags them in order that their location can be continuously monitored. Unlike most other available systems, our system focuses on using the range or distance information since they indicate how far (in terms of meters) a human target is located away from the camera and more importantly because they are capable of generating a 3D perspective of the human target (i.e. by method of 3D point clouds). The range video sequence is obtained by using a special range image camera, which is an optical imaging system which offers real time 3D image data. Furthermore, the segmented human target is modeled by image skeletonization using a star skeleton structure [7]. This model in future research can be used in conjunction with HMM's (Hidden Markov Models) for human activity recognition. The system designed calculates the angles between different body parts to analyze the posture of a construction worker. It also incorporates the use of a particle filter [2] to trace the path of the construction worker in order to classify different work-related activities. Our system is also capable of detecting multiple people and tracking each of their paths separately in a given work environment.

Keywords: Range Image Processing, Real-Time, Safety, Detection, Tracking, Surveillance, Star-Skeleton Model, Construction Work Zone.

1. Introduction

In the field of civil engineering, safety is becoming a growing concern on account of the numerous accidents occurring in construction work-sites. As per the Census of Fatal Occupational Injuries released by the U.S Department of Labor [13], out of the total of 5,448 fatal injuries which occurred in the year of 2007, 1,152 accidents occurred in the Construction industry while 1,423 occurred in the Transportation industry. This data clearly signifies the importance of increasing safety and detecting and avoiding fatal accidents before they occur in work-sites. In addition to safety, the health of construction workers is also an important issue, owing to the physical stress they are subject to while handling heavy equipment. Head injuries, spinal injuries, slip-discs, dislocations and fractures are some of the most common injuries which occur in construction sites. Most of these injuries can often result chronic health problems or in some cases even death if overlooked at construction sites.

With the help of computer vision technology we propose to device a real-time monitoring system to assist in providing safety of construction workers and monitoring the their posture while working in order to determine whether they are subject to physical health related issues due to an incorrect posture. Managing safety in construction work-sites often require an in-depth analysis of real time visual data. Although, this can be done with the help of a human operator, the chances of error and wrong analysis increases manifold as the number of workers and the area of the site increases. With the advent of real-time intelligent systems, the complexity of analyzing safety in construction sites can greatly be reduced. It is important to note that the development of real-time monitoring systems is not to replace human operators (i.e. safety officers) but instead to assist them to manage the safety of workers more efficiently and quickly with a lower rate of error especially in large construction sites with many workers performing various tasks simultaneously.

The system proposed is capable of detecting, tracking and monitoring construction workers using 3D Range Imaging technology [1]. The proposed system performs a series of steps which include background modeling, background subtraction, automatic threshold selection, noise removal [denoising], contour formation, centroid detection and tracking a human target using a particle filter [2]. The system then proceeds to form a star skeleton model which can be used to analyze the motion and posture of construction workers in a work zone site. The angles between the different segments of the star skeleton model are calculated for the purpose of motion analysis in order to detect the activity being performed by the construction worker.

2. Background Review and Range Imaging Technology

The primary purposes of tracking algorithms are to automate the process of segmenting and tagging objects from real-time video sequences. The objects which are to be tracked span a wide range including humans, vehicles, etc. Along with segmenting and tagging objects, many of the new generation tracking algorithms are also capable of analyzing the path traced by the detected objects. Our focus is mainly on developing and using tracking algorithms in civil engineering for monitoring the location of construction workers and interpreting the different activities performed them in active work zones for safety and health related issues.

Fujiyoshi and Lipton [7] developed a star-skeleton model in order to represent the human body structure. The star-skeleton model technique involves extracting the contour of the segmented object and then connecting the centroid of target object to extremities of the contour. Using this model they analyzed two motion cues namely cyclic motion and posture of the human star-skeleton model. Haritaoglu and Harvood [9] presented a method which used both shape analysis and tracking to locate humans and their body parts (head, hands, feet, and torso) and also created corresponding models in order so that they could be tracked even in presence of occlusions. The shape model used for their algorithm was the Cardboard Model [9], which represents the relative positions and size of body parts. However it is important to note that the cardboard model can be used only for upright people [straight posture]. Pfinder [3] is a real-time system for tracking a person which uses a multi-class statistical model of color and shape to segment a person from a background scene. It finds and tracks people's head and hands under a wide range of viewing condition. Yamota and Ohya [5] presented a human action recognition method based on a hidden Markov model (HMM). It is a feature-based bottom-up approach that is characterized by its learning capability and time-scale invariability. Yu and Aggarwal [8, 11] also incorporated the use of HMMs in their algorithm for detection of fence climbing from monocular video. To analyze the resulting time series, they built a block based discrete Hidden Markov Model (HMM) with predefined action classes {walk, climb up, cross over, drop down} as the state blocks. The detection was achieved by decoding the state sequence of the block based HMM. Many tracking algorithms try to solve problem in video analysis to track moving objects during a video sequence in presence of occlusions. Conte and Foggia [4] used a method of tracking objects through occlusions that exploits the wealth of information due to the spatial coherence between pixels, using a graph-based, multi-resolution representation of the moving regions. Bregler [12] uses many levels of representation based on mixture models, EM, and recursive Kalman and Markov estimation to learn and recognize human dynamics.

Range Imaging technology brings together a combination of obtaining the amplitude, intensity and the range (i.e. distance) information at every pixel of a two-dimensional sensor array of a range camera. Range Image cameras make use of ASP (Active Sensor Pixels) data of entire scene field-of-view in one frame. Range Image cameras operate on the TOF (Time of Flight) principle using phase shift measurement [1]. The camera uses conventional light emitting diodes (LED’s) to actively illuminate a scene by emitting sinusoidal modulated (spatial or temporal) near-infrared light. The time the light needs to get to and to return from an impinged object in a scene back to the sensor is then measured using a practical synchronous sinusoidal demodulation. Focused through a lens and within the 3D range camera, a CMOS/CCD sensor chip is positioned to receive the incoming wave front. It then calculates the amplitude, intensity, and range values based on the TOF principle. Both the detection and the complete demodulation are performed in the charge-domain using charge-coupled devices (CCD). This ensures almost noise free demodulation of the incoming light signal.

3. Methodology and Algorithm Design

The overall system architecture consists of seven modules namely Background Subtraction, Threshold Selection, De-Noising & Contour Formation, Centroid Locating, Tracking and Star-Skeleton Modeling [7] and Motion Analysis. The algorithm goes through five different phases which include Pre-processing, Denoising, Plotting, Tracking and Feature Recognition. The system initiates the algorithm by recording the range video sequence using a 3D range image camera. The raw data obtained by the camera is then processed to extract only the range data, while the intensity and amplitude information are discarded. Since the system works on the principle of static background modeling, initially the static background (200 frames at a rate of 25 frames/sec) of the indoor environment for the experiment is recorded. The system architecture along with a description of each of its modules involved in the algorithm is detailed below:

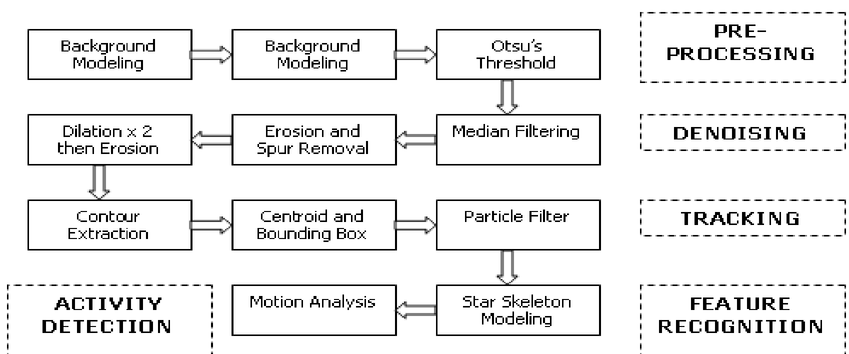


Figure 1: System Architecture

3.1 Background Subtraction

In order to extract a human target from the range video sequence the algorithm uses the method of average background subtraction. For this purpose it initially records the range video sequence of only the background without any moving human target. It then models the background as a single frame, by calculating the time average of the background (i.e. the average value of the background over all the frames). This single modeled background is then treated as the reference background frame. The range video sequences for different experiments are then recorded. Segmentation of the human target is done by subtracting the modeled background from each frame in the range video sequence, and its absolute value is then taken. Thus the system then has an array containing the absolute difference between each frame and the background of the test video sequence.

3.2 Threshold Selection

After obtaining the background subtracted image of the human target the algorithm then proceeds to threshold selection. In order to separate the foreground pixels from the background pixel (i.e. segmentation), it needs to threshold the absolute difference array. For this purpose, it uses *Nobuyuki Otsu's* method [6] of threshold selection from gray-level histograms. This method adaptively calculates the threshold based on the histogram of the difference array. The procedure utilizes only the zeroth and the first order cumulative moments of the gray level histogram to calculate the threshold. Note that the algorithm was used to calculate the threshold for each difference array of the test video sequence. Optimal sets of threshold were selected automatically and stably not based on the differentiation (i.e. a local property such as valley), but on the integration (i.e. a global property) of the histogram.

3.3 Denoising and Contour Formation

Thresholding alone, however, is not sufficient to obtain clear foreground regions; it results in a significant level of noise, for example, due to illumination changes. For proper segmentation it is important to eliminate this noise. For this purpose a series of post-processing noise removal techniques are applied on the thresholded array. Initially the thresholded array was subjected to median filtering. This was done in order to eliminate the salt-and-pepper noise present in the array. The array was then converted to a binary array (white representing the foreground pixels and black representing the background pixels). Erosion was then applied to the binary array. Erosion is one of the two basic operators in the area of mathematical morphology, the other being dilation. It is typically applied to binary arrays, but there are versions that work on grayscale arrays. The basic effect of the operator on a binary image is to erode away the boundaries of regions of foreground pixels (i.e. white pixels, typically). Then a fast binary connected component operator is applied to find the foreground regions and smaller regions are eliminated [9]. The largest blob (i.e. component) is then selected thus achieving 100% elimination of the noise pixels. After this process, all the small holes in the foreground region are filled. The position of the centroid and the dimensions of the bounding box are then calculated for the foreground regions (i.e. the segmented blob/human target). Boundary detection was applied on the segmented foreground pixels using *Canny-Edge* detection. This gives the resulting contour or boundary of the segmented blob/human target.

3.4 Centroid Location

Once the contour is formed, in order to achieve real time tracking of the blob/ human target in the range video sequence, the first step is to locate the centroid of the blob/ human target. The centroid is used to represent the human target moving in the range video sequence. The centroid calculated specifies the center of mass of the blob/human target. The centroid of the human target extracted is calculated by using the co-ordinates of the pixels on its boundary. Note that the first element of centroid is the horizontal coordinate (or x-coordinate) of the center of mass, and the second element is the vertical coordinate (or y-coordinate). The centroids are calculated for each person(s) in the range video sequence and are used as an input to the *Particle Filter*. This process serves as a primary step for tracking.

3.5 Tracking [Particle Filter]

In order to achieve tracking of the people in the range video sequence a particle filter was used. The basic idea of a particle filter [2] is to use a number of independent random variables called particles, sampled directly from the state space. In order to represent the posterior probability, and update the posterior by involving the new observations; the *particle system* is properly located, weighted, and propagated recursively according to the Bayesian rule. They are usually used to estimate Bayesian models and are the sequential ('on-line') analogue of Markov chain Monte Carlo (MCMC) batch methods and are often similar to importance sampling methods. If well designed, particle filters can be much faster than MCMC. They are often an alternative to the Extended Kalman filter (EKF) or Unscented Kalman filter (UKF) with the advantage that, with sufficient samples, they approach the Bayesian optimal estimate, so they can be made more accurate than the EKF or UKF. The centroids of the human target are extracted from each frame (using the process outlined in 5.4). The sequences of centroids over the entire sequence are then given as an input to the particle filter. The particle filter analyzes the shift in the position of the centroids over the entire video sequences and then accordingly tags and associates each blob/human target with a label. In this manner the algorithm can track when (i.e. at which frame sequence number) the human target enters and exits the range video sequence.

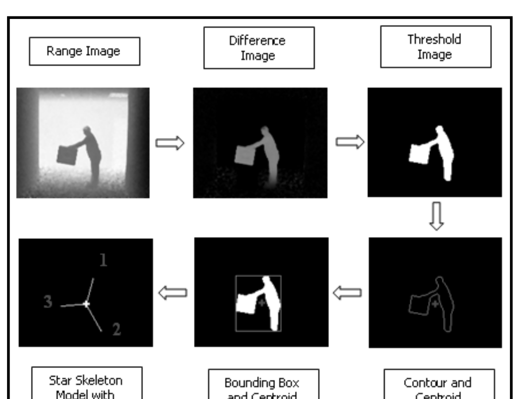


Figure 2: Algorithm Overview

3.6 Tracking Multiple People

The system designed has a provision for tracking multiple people in a given environment. For this purpose, the algorithm starts by segmenting the human targets from the range video sequence. The centroids are then detected by the algorithm for each of the human targets separately and later processed by the particle filter. The particle filter then uniquely tags and identifies each human target separately and traces its path of motion in the given range video sequence.

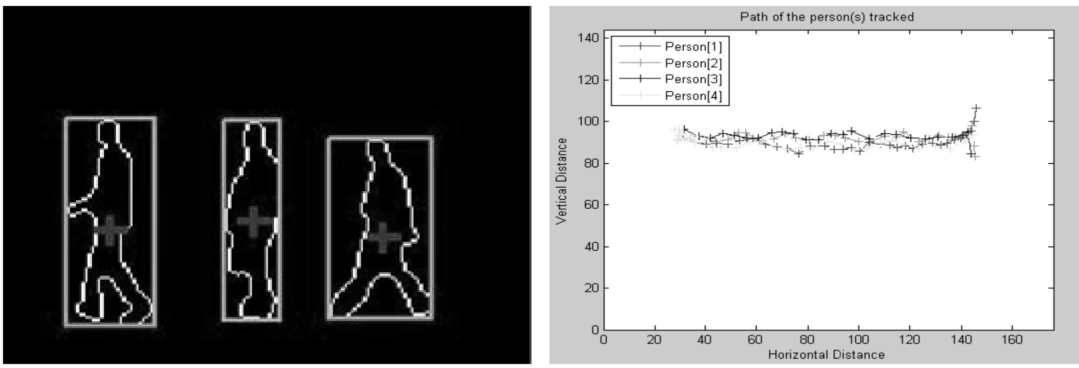


Figure 3: Particle filter output used for tracking four construction workers

3.7 Star Skeleton Modeling

In order to analyze the motion and activity of the human in the range video sequence, it is essential to extract and properly identify the extreme points lying on its boundary/contour. For this purpose the system uses the star skeleton model as proposed by Fujiyoshi. In his paper, Fujiyoshi describes a process of obtaining the star skeleton model from the contour of an extracted object, called as image skeletonization. The method proposed provides a simple, real-time, robust way of detecting extreme points on the boundary of the target to produce a star skeleton. The star skeleton consists of only the gross extremities of the target joined to its centroid in a star fashion. In order to obtain the star skeleton model the sequence of steps outlined below are followed [10]:

1. The centroid (xc, yc) of the human target's boundary is determined using the points on the contour.

2. After this, the distance (di) from each point on the boundary (xi, yi) to the centroid (xc, yc) is determined. 3. Next, low pass filtering of the signal (di) in the frequency domain is performed to remove the noise present in the signal, becoming (dic). Note that the signal d(i) is periodic with period Nb (number of pixels on the boundary of the human target) 4. The next step is to pick the maxima or peak points of the filtered signal. These points are then connected to the centroid of the target to form the star skeleton model.

The number of features (i.e. extreme points) to be detected can be controlled by the cut-off frequency (c) of the low pass filter. The higher the value chosen for (c), more extreme points can be detected in the filtered signal producing a more detailed star skeleton model. By default the value of c is kept as: $c = 0.025 \times Nb$. The main advantages of the star skeleton model are that they are computationally cheap, the scale sensitivity can be controlled and it does not depend on any apriori human model.

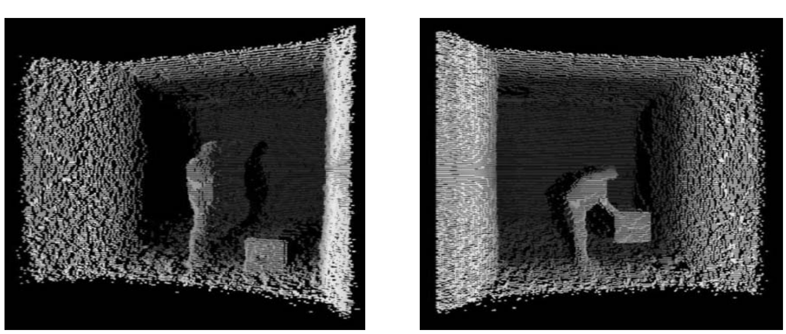


Figure 4: 3D Range Point Clouds of person picking a box

3.8 Angle Detection

Once the star-skeleton model is designed, angle detection between the various parts of the star-skeleton model is performed. The angles are calculated between the head with the left arm, the left arm with the left leg, the left leg with the right leg and the right leg with the right arm. The angle calculation is then repeated for each frame in the entire range video sequence. The angles calculated are stored in a data-structure for motion analysis. Motion analysis is a technique which is used to determine the type of motion activity which is performed by an object by analyzing the variation in the change of its path traced over a sequence of frames. The particle filter is then used, to trace the motion path of the person. However to get a more detailed view of the activity being performed by the worker, we incorporated the use of angles for motion analysis.

4. Experiments

The tracking algorithm was run on six different types of experiments which involved a worker performing various tasks in an indoor environment. The experiments included lifting a box, waving a flag to indicate warning, crawling, side-walking, sit-ups and normal straight walking. For each of these experiments, five different sets of range video sequences were recorded using a 3D range image camera. All of these experiments were recorded at a rate of 25 frames/sec. The pixel array resolution of the range image camera was 176 x 144 pixels. The output of the range camera contained the intensity, amplitude and range information of each of the pixels on the camera's sensor (176 x 144) on a per frame basis. The experiments were taken in an indoor environment due to the maximum distance constraints of the range camera.

5. Results and Analysis

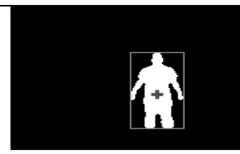
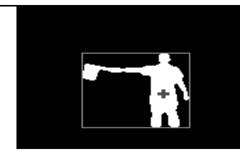
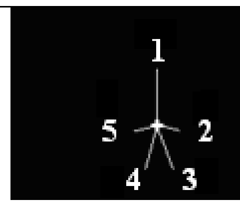
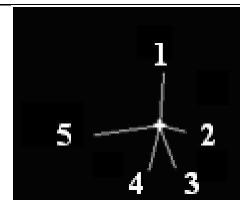
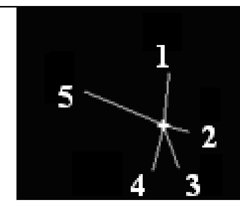
From the results in Section 7, we analyze the output of our algorithm for different test cases. Experiment 1 involves a worker performing a sidewalk experiment, in which the worker moves from left to right and vice versa. Experiment 2 involves a worker performing a crawling experiment on the ground again from left to right and vice versa. As seen above the worker has clearly been segmented from the scene and all associated noise has been cleaned from the input. The bounding box and centroid superimposed on the video sequence can be used to detect the real-time position of the worker. The star-skeleton model is in perfect conjunction with the position of the body parts of the worker as seen in the results. Once the star-skeleton model and centroid position of the worker is determined, then the motion path is traced with the particle filter designed. The plots of the tracked workers indicate the direction of motion as well as the change in the centroid position of the moving worker.

Experiment 3 involves a worker waving a flag up and down to indicate warning in a construction site. As seen in the results section, each of the body parts of the star-skeleton model are labeled successfully by our algorithm, starting with the head and then moving clockwise. From the calculations we observe that for this particular test case, the angles between [1-2], [2-3], [3-4] remain almost constant with a little variation. This is in accordance with our experiment since the relative position of the worker's head with the left arm, the left arm with the left leg and the left leg with the right leg remain almost constant in this experiment. The plot of the variation of angle between the different segments of the star-skeleton model is as shown in figure 5.

Experiment 4 taken involved a worker lifting a box. In this particular experiment all the angles between the segments of the star skeleton model vary as the worker bend down to life the box. Based on the angle of bending we can determine whether the worker is bending down in the right posture while lifting the box. The plot of the variation of angle between the different segments of the star-skeleton model is as shown in figure 6.

In experiments 3 and 4, we use the angle variations to determine if the worker is performing the activity in a correct posture. In experiment 3, where the worker is waving the flag to indicate a warning, the angle between the leg and the hand needs to reach at least 90 degrees to indicate a warning message. It needs to be observed that arm waving the flag is properly outstretched without bending. In case of experiment 4, if there is no angle between the back of the knees while bending down to lift the box, then it indicates incorrect posture. Another important observation to check is whether all the angles are varying while bending down. This is done to verify that the worker bends his back properly and does not maintain a straight back posture while picking the box

Table 1: Worker Waving a Flag

FRAME NO.	Frame 1	Frame 27	Frame 51
DENOISING AND CONDITIONING			
BOUNDING BOX AND CENTROID			
STAR SKELETON MODELING			
ANGLE DETECTION HEAD: 1 LEFT ARM : 2 LEFT LEG: 3 RIGHT LEG: 4 RIGHT ARM: 5	[4-5] = 61.05 [5-1] = 103.24 [1-2] = 104.03 [2-3] = 55.40 [3-4] = 36.26	[4-5] = 68.24 [5-1] = 102.52 [1-2] = 100.34 [2-3] = 54.09 [3-4] = 34.78	[4-5] = 99.33 [5-1] = 72.89 [1-2] = 98.88 [2-3] = 54.09 [3-4] = 34.78

We extended the angle calculation section of our algorithm to various other experiments performed and different trends were obtained for each of these cases. For example, in case of the sidewalk experiment, it was observed that the angle between the head and both arms remain constant, while the angles between the legs vary reach a minimum value of 0 degrees (i.e. when the worker joins his legs). For each experiment scenarios (walking, running, side-walking, crawling, etc.) we modeled the algorithm to analyze the angles and accordingly detected the corresponding activity performed by the worker.

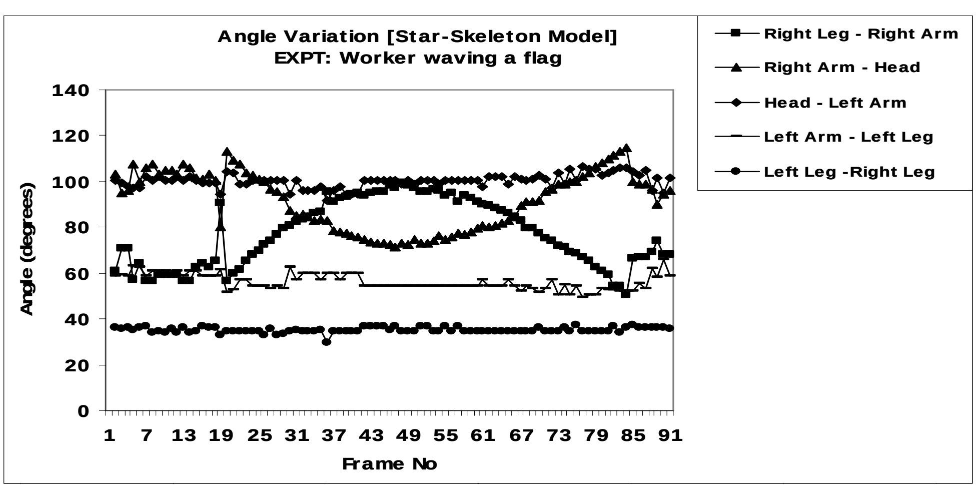
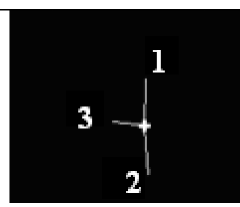
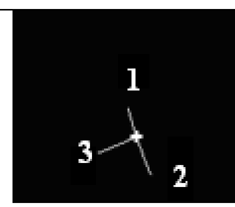
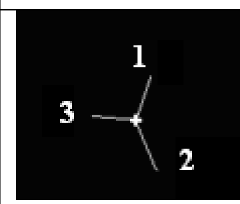


Figure 5: Plot of angle variation of the star-skeleton model for the flag experiment

The plot of the centroid variation with respect to the frame no. is as shown above for the case of the experiment of a construction worker lifting a box. From the graph we note that the worker performs the task in a periodic manner (i.e. same posture while bending each time). We also observe that the construction worker does not bend properly since the minimum distance of the centroid from the ground is beyond the defined threshold value. If the construction worker had maintained a correct posture while bending down, then the centroid of the tracked target would have reached closer to the ground while lifting the box. This condition was not observed which indicates an incorrect posture. In a similar way, for the remaining experiments by setting a threshold parameter (e.g. angle between legs, angle between right arm and head, maximum centroid, etc), we can determine if the construction worker is maintaining a correct posture to avoid chronic physical stress related conditions.

Table 2: Worker Lifting a Box

FRAME NO.	Frame 23	Frame 35	Frame 60
DENOISING AND CONDITIONING			
BOUNDING BOX AND CENTROID			
STAR SKELETON MODELING			
ANGLE DETECTION HEAD: 1 LEFT LEG: 2 RIGHT LEG: 2 LEFT ARM : 3 RIGHT ARM: 3	[1-2] = 173.59 [2-3] = 84.20 [3-1] = 102.19	[1-2] = 174.49 [2-3] = 96.53 [3-1] = 88.96	[1-2] = 138.73 [2-3] = 103.24 [3-1] = 118.02

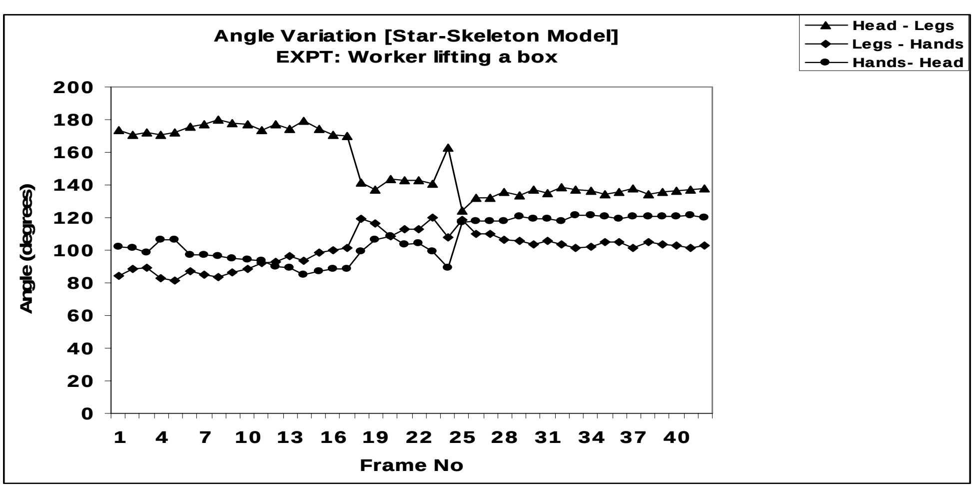


Figure 6: Plot of angle variation of the star-skeleton model for the box experiment

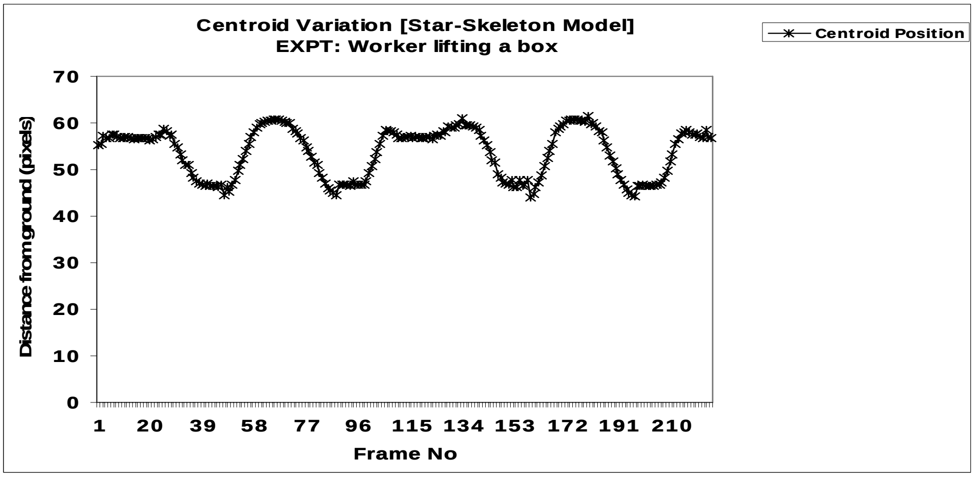


Figure 7: Plot of centroid variation of the star-skeleton model for the box experiment

6. Conclusions and Future Work

We present a real-time system for the tracking and surveillance of construction workers in a work-site for safety and health monitoring purposes. The safety of the worker is monitored continuously by tagging the target using a particle filter and then following it in the work-site. For extending the use of our system for health related issues, we modeled the worker using a star skeleton structure and then performed motion analysis to determine the variation in angles between the various segments of the model. On basis of this variation we can determine if the worker has performed the task maintaining a correct posture. Our system designed uses a 3D range image camera to segment a given target from a scene using distance based information thus providing segmentation in 3D space. We also were able to detect multiple workers and track and monitor each of them independently in the work-site. The algorithm is computationally inexpensive on resources making it a fast and robust real-time tracking system. Our system if used in conjunction with neural network training algorithms would certainly help to avoid fatal accidents as well as physical health related issues which are an important concern in the construction and transportation industry. Due to the maximum distance constraint imposed by the 3D range image camera we limited the scope of our experiments to an indoor work environment. In future we will extend our system to outdoor environments where several issues which include dynamic change in the environment and higher levels of noise will occur. Also we also plan to incorporate the use of an automated neural network system for making the training process of the system more robust.

References

[1] Teizer, Jochen, 3D range imaging camera sensing for active safety in construction, ITcon Vol. 13, Special Issue Sensors in Construction and Infrastructure Management , pg. 103-117, 2008.

[2] J. C. Crocker and D. G. Grier, "Methods of digital video microscopy for colloidal studies," J. Colloid Interface Sci. 179, 298-310, 1996.

[3] C. Wren , A. Azarbayejani , T. Darrell , A. Pentland, Pfinder: real-time tracking of the human body, Proceedings of the 2nd International Conference on Automatic Face and Gesture Recognition (FG '96), p.51, Oct. 14-16, 1996.

[4] Donatello Conte , Pasquale Foggia , Jean-Michel Jolion , Mario Vento, A graph-based, multi-resolution algorithm for tracking objects in presence of occlusions, Pattern Recognition, v.39 n.4, p.562-572, April, 2006.

[5] J. Yamato, J. Ohya and K. Ishii. "Recognizing Human Action in Time-Sequential Images using Hidden Markov Model," Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, pp. 379--385, 1992.

[6] N. Otsu, "A Threshold Selection Method from Gray Level Histograms," IEEE Trans. Systems, Man, and Cybernetics, vol. 9, pp. 62-66, 1979.

[7] Hironobu Fujiyoshi, Alan J. Lipton, "Real-time Human Motion Analysis by Image Skeletonization," wacv, p.15, Fourth IEEE Workshop on Applications of Computer Vision (WACV'98), 1998.
[8] Elden Yu, J.K.Aggarwal, "Detection of Fence Climbing from Monocular Video," icpr, pp. 375-378, 18th International Conference on Pattern Recognition (ICPR'06) Volume 1, 2006
[9] I. Haritaoglu, D. Harwood, L.S. Davis, "W4: A Real Time System for Detecting and Tracking People," cvpr, p.962, 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition.
[10] Hsuan-Sheng Chen , Hua-Tsung Chen , Yi-Wen Chen , Suh-Yin Lee, Human action recognition using star skeleton, Proceedings of the 4th ACM international workshop on Video surveillance and sensor networks, October 27-27, 2006, Santa Barbara, California, USA.
[11] Elden Yu , J. K. Aggarwal, Detection of stable contacts for human motion analysis, Proceedings of the 4th ACM international workshop on Video surveillance and sensor networks, October 27-27, 2006, Santa Barbara, California, USA.
[12] Christoph Bregler, Learning and Recognizing Human Dynamics in Video Sequences, Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR '97), p.568, June 17-19, 1997.
[13] U.S Department of Labor, Bureau of Labor Statistics, http://www.bls.gov/news.release/cfoi.t03.htm

Integrating Automated Data Acquisition Technologies for Progress Reporting of Construction Projects

Samir El-Omari[1] and Osama Moselhi[2]

1 P. Eng., Assistant Professor, Department of Building, Civil and Environmental Engineering, Concordia University, 1455 De Maisonneuve Blvd. W., Montreal, Quebec H3G 1M8, Canada, s_elomar@encs.concordia.ca
2 P. Eng., Professor, Department of Building, Civil and Environmental Engineering, Concordia University, 1455 De Maisonneuve Blvd. W., Montreal, Quebec H3G 1M8, Canada, moselhi@encs.concordia.ca

Abstract

Controlling construction projects necessitates controlling their time and cost in an effort to meet the planned targets. Management needs timely data that represent the status of the project to take corrective actions, if needed. This paper presents a control model that integrates different automated data acquisition technology to collect data from construction sites required for progress measurement purposes. Current automated data acquisition technologies are described, and their suitability for use in tracking and controlling construction activities is assessed. This includes bar coding, Radio Frequency Identification (RFID) 3D laser scanning, photogrammetry, multimedia, and pen-based computers. The user can move with a tablet PC in the construction site and record, take snapshots and also hand written comments about activities on site. The proposed cost/schedule control model Integrates with the automated data acquisition technologies, a planning and scheduling software system, a relational database, and AutoCAD to generate progress reports that can assist project management teams in decision making.

Keywords: 3D laser scanning, photogrammetry, RFID, Tablet PC, Bar Coding progress reporting, data acquisition, automation

1. Introduction

The earned value technique is widely used for periodic monitoring of actual expenditures and physical scope accomplishment and, accordingly, for generating period-by-period progress reports. These reports are commonly developed by essentially comparing the collected actual data pertinent to work performed on site to that planned. The reliability of these progress reports depends primarily on the accuracy, and timeliness collection of actual data that depicts work progress on site. This paper presents a control model that integrates different automated data acquisition technologies including bar coding, Radio Frequency Identification (RFID) 3D laser scanning, photogrammetry, multimedia, and pen-based computers to collect actual data from construction sites to generate progress reports. To do so, the characteristics of different automated data acquisition technologies were studied and analyzed. This includes their capabilities and limitations and their respective suitability to track various construction operations. Experiments were conducted to study the applications of different automated data acquisition technologies and explore the most suitable IT platform for integrating them in one tracking and control system Each automated technology, is used for a certain construction task on site. For example, 3D scanner or LADAR (laser distance and ranging) was integrated together with photogrammetry to rapidly track changes of quantities of work accomplished such as excavation works. Integrating these two technologies alleviates limitations associated with each of them individually such as the number of scans required and the time needed for each scan to produce acceptable results during the 3D modeling process. It also overcomes limitations associated with photogrammetry when modeling 3D images of objects with unclear geometrical properties as in the case of earthmoving operations where modeling 3D images from digital photo images becomes difficult and the presence of a scanned image can be helpful. Bar coding and RFID are utilized for material and labor tracking. In the reporting stage, more photo images would be more desirable. Pen-based or tablet computer is utilized as the main interface tool with the user [1].

2. Proposed Control Model

A project control system establishes guidelines for effective cost and schedule control. As mentioned earlier, data collection is a crucial step in the tracking control process. Considerable work has been carried out to utilize various automated data acquisition technologies for the purpose of data collection [2], [3], [4], [5], [6]. For example in Abudayyeh's model [7], the bar code technology was used in acquiring construction data from site. In some other cases, these technologies were used for the purpose of inventory or inspection. The proposed model integrates different automated data acquisition technologies. Construction data is collected from site and stored into a centralized database for later use in generating progress reports. The system is designed from the contractors' point of view to help track their projects in a timely manner. It also allows owners to have a closer look over the project. Pen-based or tablet computer is utilized as a media of integration. An interface software application was developed using Microsoft Access. With this application, actual data related to each activity on the job site is collected using a pen-based computer, RFID, bar coding, LADAR, or multimedia information in form of voice records or video tapes.

3. Study and Evaluation of Site Data Acquisition Technologies

This research study involved experimenting with different hardware and software systems. This section presents the work conducted on laser scanning, photogrammetry, and RFID to find the best combination that would provide the desired results. Comparisons with other available tested components are presented and the reasons behind the selection of the components used in the developed system are highlighted. Table 1 provides a comparison between the proposed model and other available integration models and systems. The comparison was carried out with respect to the purpose and essential characteristics of each model and system Included in the comparison are the Building Information Modeling (BIM) [8], Documentum by EMC Corporation [9], Coreworx from Software Innovation Inc. [10], and PM+ of SNC-Lavalin [11].

BIM includes geometry, spatial relationships, geographic information, quantities and properties of building components. It extracts the information from design drawings and models it in 3D images to represent the site conditions and also calculate quantities of work performed. Future research could integrate features of the proposed system with BIM in regard to collecting actual information from the construction site [12], [13], [14]. Documentum is document management software and it manages document content and attributes such as check-in, check-out, and version management. The purpose of proposed system isn't document control, but to assist in collecting and processing near real time data from construction sites such as labor or equipment hours to update the project schedule and generate progress reports. Coreworx is similar to Documentum, it automates document control functions, such as, version control, approval workflows, transmittal receipt/generation and built-in document reporting. PM+ developed by SNC-Lavalin manages progress payments for suppliers and sub-contractors but it doesn't automate the process of data collection as in the case of the proposed control system. Data pertinent to material usage and labor and equipment hours are entered in PM+. PM+ also has a document control module but it doesn't store the documents in the software, it only provides information on the physical location of the document. From the comparison above, it can be noted that the proposed system utilizes different technologies for data collection needed to update the project schedule. This feature is not available in the other systems included in Table 1.

4. Laser Scanning

Laser scanning operation components include a scanner, which is connected to a laptop computer through a serial connector type RS-232 or RS-422 depending on the scanner type or through a TCP/IP network cable. The scanner cannot operate without a scanning software installed on the laptop and that is the reason for the presence of the laptop. The scanning software can enhance the scanned image, such as removing destructing points caused by some obstacle from the point cloud image, before exporting it to the modeling software application, which is the last step in the 3D scanning operation. The triangulation based laser scanning technology was developed as early as 1978 and the National Research Council of Canada was among the first institutes to develop it [17]. 3D scanning equipments evolve very rapidly. Table 2 provides information on some of the currently available 3D scanners. Scanner speed (how many points per second can it reads), range (distance from the scanner to the object to be scanned), and accuracy are important factors that are considered in new scanners. The scanner employed in this research was purchased in 2002 and it was one of the most advanced scanners at that time. Since then scanners have improved dramatically in their speed and accuracy but their prices remain a major limitation. The proposed control system utilized 3D scanners for progress measurement purposes. During the pace of this study, some software systems were experimented with and it was decided to integrate photogrammetry to enhance the productivity and accuracy of 3D modeling for progress measurement purposes [15], [16]. The reason for this integration is to reduce the number of scans required, as more information can be extracted from the photo images. It also helps in modeling 3-D images without the need to acquire point cloud images with high accuracy. Photo images are taken with a regular digital camera to assist in modeling 3-D images. The method is explained by the authors in [16]. Figure 1 illustrate the 3D modeling process currently undergoing to extract quantities of HVAC ducts in a building construction project.

Table 1 Comparison between the proposed model and others

	Proposed model	BIM	Documentum	Coreworx	PM+
Main purpose	Integrating different data acquisition technologies to collect data from construction sites needed for progress measurements	3D modeling from AutoCAD drawings for site representation and quantity measurements and help in developing as-built information	Primarily Document management	Primarily document management	Cost estimating, document control, and progress measurement of work packages
Level of integration	Activity level	3D models that represent building components (e.g., beams, columns, walls) or assembly (e.g., floors, stairwells, facades)			Work package level
Document control	No	No	Yes	Yes	Yes, but doesn't store documents
Laser scanning	Yes, used for calculating actual quantities and site representation	No	No	No	No
Photogrammetry	Yes, used for calculating actual quantities and site representation	No	No	No	No
RFID / barcode	Yes, used to collect actual working hours	No	No	No	No
Modeling 3D from AutoCAD	No, AutoCAD is integrated but only for information purposes	Yes	No	No	No
Modeling from scanned and digital images	Yes	No	No	No	No
Site representation from design drawings	No	Yes	No	No	No
Integrating planning software	Yes, to import and update planned data	No	No	No	Yes
Automated collection actual working hours	Yes	No	No	No	No
Calculating quantities from actual work	Yes	Quantities are calculated from the design drawings	No	No	No

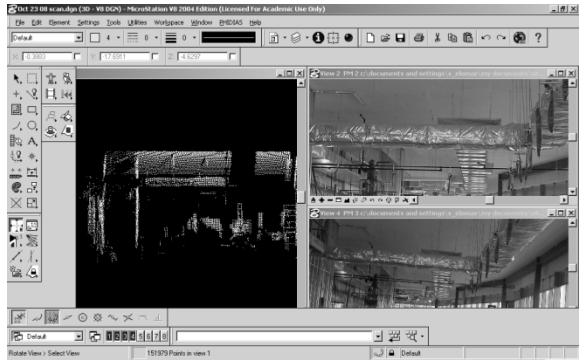


Figure 1 3D modeling from scanned and digital images

Table 2 Laser scanners comparison

	Company	Equipment	Range	Accuracy	speed	Price	comments
			m	mm	Points / Sec	1000 $US	
1	Steinbichler Optotechnik	T-Scan 2	0.08	0.03	10,000	150 - 250	Hand held
2	Leica Geosystems	Leicascanstation 2	300	6	50,000	120	
3	Leica Geosystems	Leica HDS 6000	79	6	500,000	120	
4	FARO Technologies	Faro Laser Scanner LS 840/880	79	3 at 25 m	120,000	100	
5	Datapixel	Optiscan H class	0.1	0.006	60,000	100	including arm
6	3D Digital	Optix 400	0.3 - 0.9	0.0035	1,000,000	45	
7	3rdTech	Deltasphare-3000	15	7	24,000	30	
8	CALLIDUS Systems	CALLIDUS CP 3200	80	5	1,750	120	
9	RIEGL Systems	LPM-100VHS	200	2	1,000	85	
10	RIEGL Systems	LPM 321	6000	25	1,000	131	
11	RIEGL Systems	LMS-Z420i	1000	10	11,000	149	
12	RIEGL Systems	LMS-Z390i	400	6	11,000	125	

5. Bar Coding and RFID

The application of bar coding and RFID required thorough investigation of the available equipments in the market. RFID readers are either fixed readers, which are known as stand-alone readers, hand held readers and vehicle mounted readers. All three types of readers can be utilized for progress measurement through reading RFID tags of resources to update their working hours. The RFID technology grows fast and will soon replace bar-code technology when the price of utilizing it competes with bar coding technology.

A Hand-held RFID reader was purchased for this research because it can read bar-code and RFID tags. The reader IP4 is from Intermec Inc. [18], [19] it combines a RFID reader and the 751A mobile computer, which makes it capable of reading RFID and bar-code tags as well. The operating system for 751A is Microsoft Windows for pocket PC. Intermec readers use the basic reader interface (BRI) command-language that enable to read RFID tags and write information on them. The users must develop a user interface designed for their needs with programming languages like C$^{\#}$, C/C^{++}, .NET or Java and use BRI command-language for tasks to be performed by the RFID reader. Figures 1 to 5 illustrate a demo application for reading and writing on RFID tags using the BRI command language. Intermec have recently developed mobile computer (CNe and CNe3) [19] with GPS capability that can be attached to the new IP30 RFID reader. GPS RFID tags are now available in the market to easily identify the location of the tag (IDENTEC SOLUTIONS) [20].

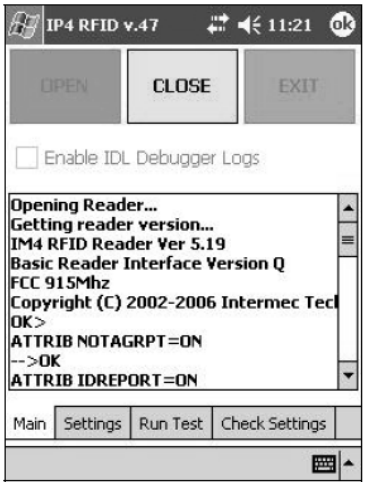


Figure 2 BRI demo from Intermec start screen

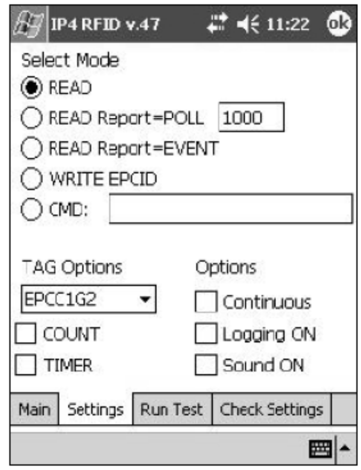


Figure 3 BRI demo setting screen

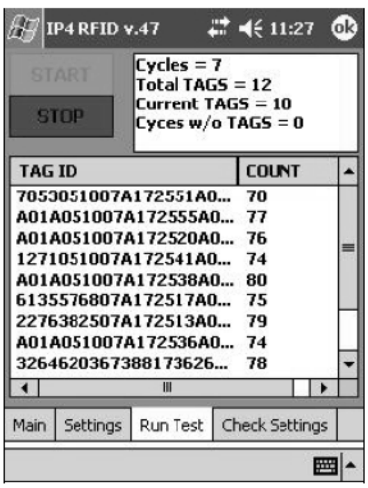


Figure 4 BRI demo RFID tag reading

Figure 5 BRI demo write on RFID tags

6. Components and Data Flow of the Developed Model

The system integrates through its pen-based computer environment, RFID, bar coding, LADAR and multimedia technologies to automate data acquisition. At its core is a relational database, used to organize and store the collected data. An interface software application is developed in the pen-based computer environment (Fujitsu - Stylistic ST4121B). The developed interface integrates different construction management software applications such as Primavera Project Planner, Microsoft project, AutoCAD, Microstation and other software applications such as PHIDIAS, which is capable of integrating and modeling photo images together with scanned images acquired using the LADAR (LPM 100 VHS). In the proposed system, scanned images are modeled with the support of digital photo images to produce 3D images of the scanned object, which are then used to estimate quantities of work performed. The layout of the proposed control system is shown in Figure 6 [1], [21]. Data pertinent to equipment working hours, labor hours and material consumption are collected using the developed system. Scanned and digital images are modeled to determine quantities of work performed and subsequently calculate percent of work complete. Percent (%) complete can also be determined through templates, which assign different weight to the sub-tasks involved in the activity being tracked [22]. During the process of data acquisition, the user can access AutoCAD drawings related to the activity in-progress and record notes for any changes occurred to help in the preparation of as-built drawings. Information needed in the future about the performance of the project within a specified period can, accordingly, be retrieved for later use. Clearly, the EV concept works well if the data needed to generate that type of control are accurately collected in a timely manner. Traditionally, actual data pertinent to material use, man-hour, and/or equipment use, is collected manually by filling forms on site and then feeding the collected data into a computer in the office. This is not only time consuming but also is susceptible to human error, and may lack consistency and reliability. The proposed system aims at alleviating these shortcomings by incorporating the automated data acquisition technologies described in [1], [15], [16], [18], [21].

7. Proposed Model's Database

The database of the system was designed to organize and store data collected from construction sites that support the management functions of the proposed model [1]. The database is of a relational type, and consists of 37 entities or tables. The entity relationship diagram is shown in Figure 7. The database was implemented in Microsoft Access to facilitate the interaction with scheduling software systems such as Microsoft Project and Primavera systems. The main entities of the database are Projects, Activities, Labors, Equipment, Materials, Photos, Sound, Videos, 3D Images and Drawings. The attributes of these entities are listed in Figure 7. A "Time Sheet" entity was constructed to store daily start and finish time of labors and equipments. Other entities were designed such as the "Activity Drawings" entity, which holds the primary keys of the "Activities" and "Drawings" entities, to realize the many-to-may relationship between both activities. Other types of relationships that exist beside many-to-many are the one-to-many as in the case of Projects and Activities because a project can have many activities and an activity belongs only to one project. To update planned information, data had to be first imported from the output of the scheduling software such as Microsoft Project. Microsoft Project can export data in ACCESS type files. An example was performed on the JMSB project currently under construction at the University to export planned data, update it, and send it back to Microsoft Project. To update the proposed system's database with information imported from the scheduling software, a number of queries were developed using SQL language. A list of forms was designed to facilitate the interaction process with the user. The user starts by a validation process of his username and password. Upon granting access, the user is prompted with the main screen shown in (see Figure 8) that includes the list of forms available is prompt so that the user can start first by selecting the project he wants to invoke. The project form, shown in Figure 9, includes project information such as its duration, and start and finish time. A list of projects is available to choose from in the project dropdown menu. The user then can access the project's activity by selecting the activity command button in the project form or from the main screen. The control process is performed on the activity level where hours spent on that activity such as labor hours has to be reported. This can help in integrating cost and time reporting. Clearly, once the data becomes available at that detailed level, it could be rolled up at the cost account or work package level. The activity form, shown in Figure 10, includes, aside from the activity information, command buttons to invoke different forms and queries. This query updates the fields: "Description", "Duration", "Start Date", and "Finish Date" with those imported from the project's schedule and with the activity ID set as criteria. Organizing collected data is a very important step and it facilitates future retrieval of this data, which can help not only in progress reporting but also in management of claims and in production of as-built drawings. If for example, a note was written regarding a particular problem related to an activity that has to do with weather condition during the data collection process then, this note is stored with its ID in the "Notes" entity along with its date and time. A link was established between the "Notes" entity and the note files so as to minimize the size of the database. Similarly, links were also constructed between photograph, video clips, drawings, and 3D modeled images and their respective "Photos", "Videos", "Drawings" and "3D Images" entities. The "Quantity to Date" attribute of the "Activity" entity is to register updated quantity of work accomplished. This quantity is calculated using modeled 3D scanned images merged with digital photo images as explained earlier. The "3D image" entity includes 3D modeled images that are linked with different activities to illustrate the current status of the project. After updating data of the activities involved during a reporting period, the database of Microsoft Project is then updated using update queries in a reverse way to that performed at the beginning and exported to the project schedule. EV analyses are performed in Microsoft Project using the calculated quantities of work accomplished described above and progress reports are generated. Additional reports such as notes, and photographs are also generated using the proposed system to highlight critical problems.

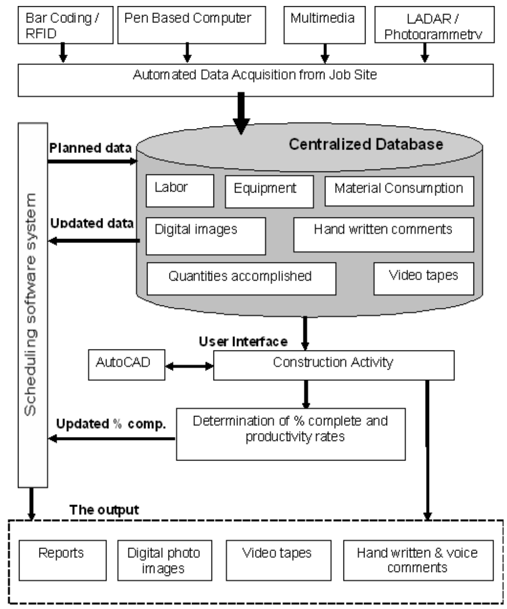


Figure 6 Proposed system architecture

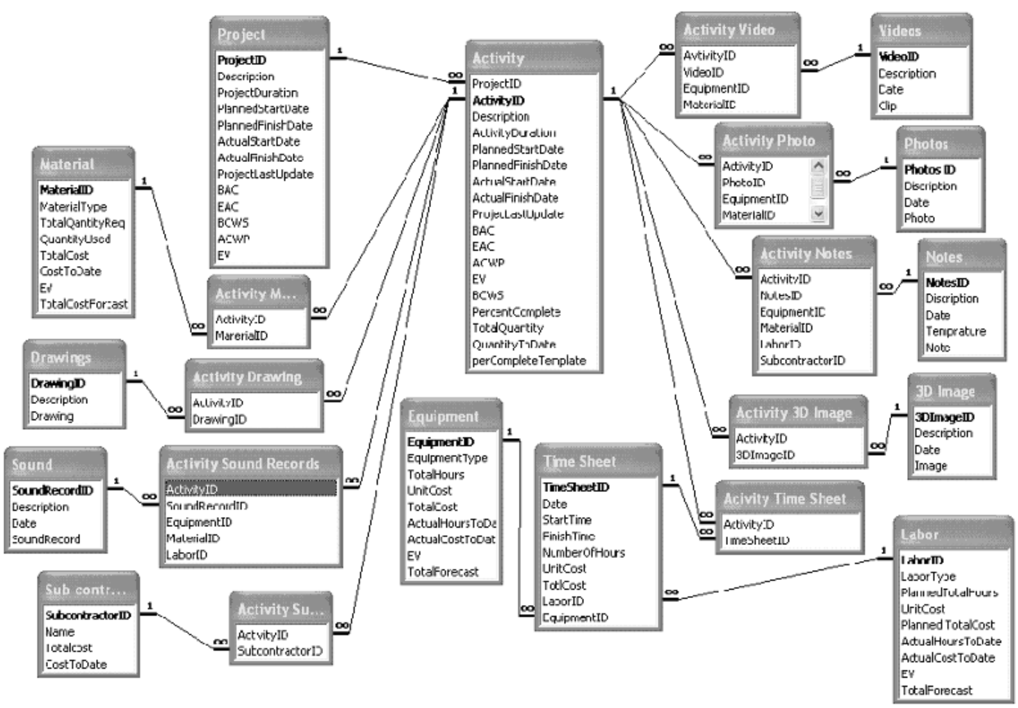


Figure 7. ER-Diagram of Developed Database

8. Conslusions

This paper presented the layout of an IT-platform, designed to facilitate automated data acquisition from construction sites to support efficient time and cost tracking and control of construction projects. The paper also described the proposed cost/schedule control system main components. A set of automated data acquisition technologies was briefly described and their potential use in construction highlighted. The system presented in the paper is capable of capturing text, numerical and graphical data to report efficiently on the project progress. Integrating Laser scanning and photogrammetry was necessary to overcome limitations associated with both technologies. The authors described this integration in [1]. Database was designed to assist management teams in performing project tracking and control functions in an efficient manner, and in management of construction claims. The proposed model's database was described and its main entities were highlighted.

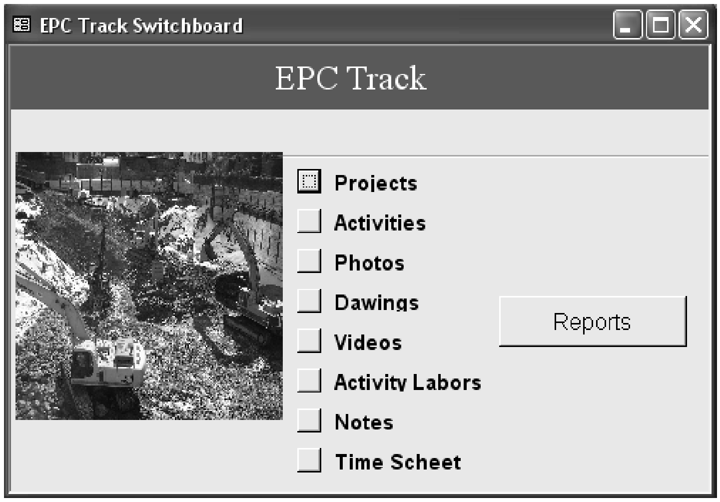


Figure 8 EPC Track switchboard

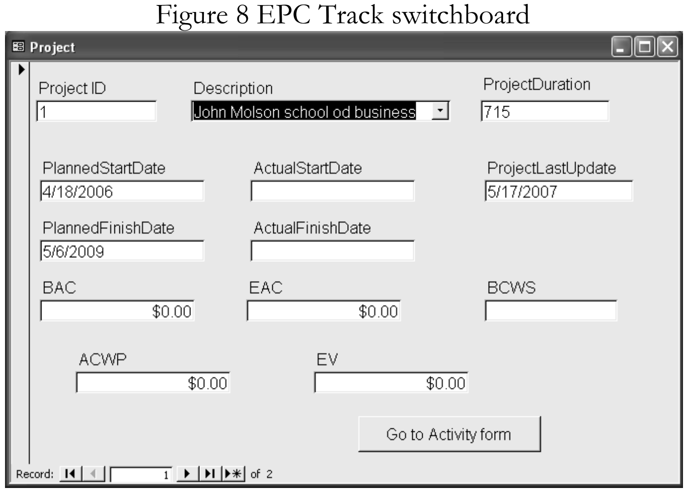


Figure 9 Projects form

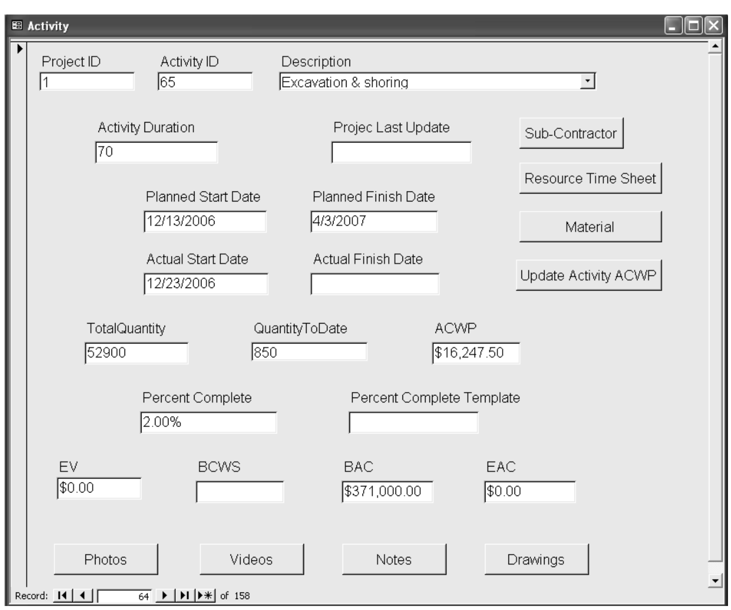


Figure 10. Activity form

Acknowledgements

The authors would like to acknowledge the financial support provided by the National Science and Engineering Research Council of Canada in the form of operating grant

References

[1] El-Omari, S., 2008. Automated Data Acquisition for Tracking and Control of Construction Projects, Ph.D. thesis presented to the department of building, civil and environmental engineering at Concordia University Montréal, Québec, Canada
[2] Akinci, B., Patton, M., Ergen, E. (2002). "Utilizing Radio Frequency Identification on Precast Concrete Components-Supplier's Perspective." Proc. Of the 19th International Symposium on Automation and Robotics in Construction., NIST, Gaithersburg, MD, September 2002.
[3] Perera, A., Imriyas, K. (2004). "An Integrated Construction project cost Information System Using MS Access and MS Project." J. of Constr. Mgmt and Economics, 22, 486-502.
[4] Shashi, M., Jain, K (2007)." Use of Photogrammetry in 3d Modeling and Visualization of Buildings." ARPN J. of eng and app sci , 2, NO. 2, APRIL 2007
[5] Du, J., C., Teng, H., C., 2007. 3D Laser Scanning and GPS Technology for Landslide Earthwork Volume Estimation. Automation in Construction, 16 (2007) 657 – 663.
[6] Jaselskis, E. (2003). "Implementing Radio Frequency Identification in the Construction Process." J. Constr. Engrg. and Mgmt, ASCE, 129(6), 680-688.
[7] Abudayyeh, O., 1991. "An Automated Data Acquisition and Data Storage Model for Improving Cost and Schedule Control," Ph.D. thesis presented to the department of civil engineering at the North Carolina State University
[8] Eastman, C., Teicholz, P., Sacks, R., Liston, K., 2008 "BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors", John Wiley & sons, 2008.
[9] EMC, http://www.emc.com/products/family/documentum-family.htm
[10] Software Innovation Inc., http://www.softinn.com/
[11] SNC-Lavalin, http://www.snc-lavalin.com/
[12] Lee, G., Sacks, R., Eastman, C. M. 2006. "Specifying parametric building object behavior (BOB) for a building information modeling system". Automation in Construction, 15(6), 758-776.
[13] Holness, G., 2008, "Building Information Modeling Gaining Momentum." ASHRAE Journal. Pp 28-40. June 2008.
[14] Eastman, C., Teicholz, P., Sacks, R., Liston, K., 2008 "BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors", John Wiley & sons, 2008.
[15] El-Omari, S., Moselhi, O., (2007). "Hybrid Methodology for Automated Collection of Data from Construction Sites" International Symposium on Automation & Robotics in Construction ISARC - 2007 Kochi, Kerala, India
[16] El-Omari, S., Moselhi, O., (2008) "Integrating 3D Laser Scanning and Photogrammetry for Progress Measurement of Construction Work", Automation in Construction, 18 (2008) 1–9
[17] Mayer, R. (1999), "Scientific Canadian: Invention and Innovation From Canada's National Research Council", Vancouver: Raincoast Books, 1999.
[18] Moselhi, O., El-Omari, S. (2006). "Integrating Bar Coding and RFID to Automate Data Collection from Construction Sites", Joint International Conference on Computing and Decision Making in Civil and Building Engineering, ASCE, June 14-16, 2006, Montreal, Canada.
[19] Intermec Inc., http://www.intermec.com/
[20] Identec Solutions, http://www.identecsolutions.com
[21] Moselhi, O., El-Omari, S., (2005), "The Use of 3D Scanners for Automated progress Reporting on Construction Activities", AACE 49thInternational Annual Meeting, 2005, New Orleans, Louisiana
[22] Moselhi, O. (1993). "Applied Earned Value for Project Control." CIB W-65, Trinidad, W. I., 869-879.

An Analytic Model Combining Monte Carlo Simulation and PSO in Estimating Project Completion Probability of Project Durations and Costs

Henyi Jen[1*] and Yu-chiun Chen[2]

[1,2]Department of Industrial Engineering and Management, Yuanze Univerisity, 135 Yuan Tung Road, Jhongli, 320, Taiwan; PH: +886-3-4638800; Email: henyi@saturn.yzu.edu.tw

Abstract

Project uncertainties are always the reason of project delay or budget overrun. Especially in tight schedule or project crashing, it is hard to balance both project duration and costs. Past research focused on the optimal schedule and costs, without knowledge of its on-time, within-budget completion risk. This research provides an analytical model by first using PSO heuristic algorithm to find the minimum project costs under time constraint. Monte Carlo Simulation is then implemented to build a completion probability table of time/cost combinations. The time and cost from PSO method is compared with the probability matrix. An analysis is provided as a demonstration.

Keywords: Time-cost trade-off problem, PSO, Monte Carlo Simulation, Probability Matrix

1. Introduction

In project management, time and cost are two control factors. Project should be delivered in time. However, project execution is often impacted by the uncertainties and thus may delay the committed project deadline. This delay not only taints the reputation of project manager, but in some cases, the project owner may claim liquidation damages from project performer. High penalty costs and loss of revenue would be ensued. Project performer usually concerns costs more than time in taking projects; nonetheless, only when both time and costs are concerned simultaneously and equally, can a project be taken with minimal chance of loss of profit.

The problem to find the optimal schedule and cost combination of a project is NP-hard. Most scholars approach this type of problems by using heuristics to find better solutions. This study approaches this problem from risk perspective by using the project completion probability of specified time and costs to see if it is feasible for a project performer to take a project. A combined Monte Carlo simulation with Particle Swarm Optimization (PSO) heuristic will be used in the study to evaluate the completion probability of a project under various time and costs combinations.

2. Previous Study

Project activity cost varies with its duration and resource inputs. On the other hand, different resource inputs and costs will vary the activity duration as well. This is categorized as Time-Cost Trade-Off Problem (TCTP). Kelly (1961) assumes the relationship between activity time and cost to be linear to reduce the difficulty in solving the TCTP. Meyer and Shaffer (1963), Patterson and Huber (1974) suggest to use mixed integer programming technique to handle the TCTPs that have discrete and linear relationship between time and costs.

Project execution is full of many uncertain factors. Examples include unforeseen bad weather that causes the interruption of the project progress; the project owner may request scope change and the change order due to unknown job site conditions. Even the labor and the equipment have inherent variability. The materials may be delivered irregularly or not sufficient materials when needed. The simplest method to process the uncertainties is to use Monte Carlo simulation (Diaz and Hadipriono, 1993). Other approaches also include probabilistic TCTP or the combination of both deterministic and probabilistic TCTP (Laslo, 2003; Ke and Liu 2005) and heuristic methods (Daisy et al. 2004, Yang 2006). Yang (2006) introduces PSO heuristic in the project crashing analysis. Total minimized project costs can be found efficiently under specific time window should the activities be crashed in three forms: discrete, piecewise-linear, and non-linear.

3. Research Methodology

The successful implementation of PSO heuristic into project crashing analysis with various types of activity encourages the authors to further the study into the analysis of optimal project time and costs by performing the combined Monte Carlo simulation and PSO heuristic method. This combined method is executed according to the following three steps.

3.1 Step 1: Use PSO to find the minimum project direct costs with specific time constraint

In the project, for any activity Ai, the direct activity cost is Ci, the objective function is to minimize the total costs of the summation of all activities' costs.

Min ΣCi (1)

subject to:

ESi+ti-ESj≦0 (2)

D=max{ESi+ti}≦ $\overline{D}$ (3)

ESi, ti>0 (4)

Ci=fi (ti) (5)

Equation (2) states that the earliest start time of activity Ai plus its time duration cannot exceed the earliest start time of its following activities. Equation (3) limits the time of last finished activity to be less than or equal to the specified date $\overline{D}$. Equation (4) states the earliest start time and the time duration for every activity should be greater than zero. Equation (5) indicates that the activity cost is the function of time. As previously mentioned, the time and cost relationship is complex and the study uses the definitions from Yang (2006), namely, the three various time and cost relations—discrete, piecewise-linear and non-linear functional form to reflect the different impact of time to the activity costs.

PSO heuristic is in play to handle the complex relation stated in equation (5), the solving procedure is as below:

1. Generate Initial Random Solution

Let xij be the position of j particle of activity Ai, where
i =1,2,…,N, and N is total number of activities,
j =1,2,…,M, and M is the number of particles

xij can be generated from uniform distribution function in the range between 0 and 1. In PSO term, the xij is encoded. Each xij stands for the time duration for Ai, or is decoded. For example, if xij=0.7, and if the upper and lower bounds for xij are 20 and 10 days respectively, xij is decoded as 10+0.7*(20-10), or 17 days. The positions of PSO particles can be decoded to the time durations of all activities. The activity costs can be calculated from equation (5). When the time and duration of each activity are decided, the project time can be derived using Critical Path Method (CPM) and its respective total costs can be summed up. This project total cost will become the initial solution.

2. Define Fitness Function

The total project cost is defined as the fitness function of PSO heuristic. Please note the cost outlay is only effective when the project duration should be less than or equal to $\overline{D}$, as stated in equation (3)

3. Update the velocity vector

PSO heuristic adjusts xij values to be more suitable through the update of velocity vectors. Following is the velocity vector and the updating mechanism.

Vid_new=ω* Vid_old+φi* (Pid-xid)+ φg* (Pgd-xid) (6)

Xid_new=Xid_old+Vid_new (7)

Vid_old=Vid_new (8)

Xid_old=Xid_new (9)

Here:

Vid_new: Velocity vector for i particle in d dimension (new)

Vid_old：Velocity vector for i particle in d dimension (old)

Xid_new：Position vector for i particle in d dimension (new)

Xid_old：Position vector for i particle in d dimension (old)

ω:inertia weight, φi and φg: parameters of weight, it is usually within [0,2]

Pid：Best fitness value of individual particle

Pgd：Best fitness among all particles

In the beginning, xij and the velocity vectors are generated randomly. Equation (6) is composed of three parts: the inertia of particle's previous behavior, as indicated by inertia weight ω; the cognitive consistency part, as indicated by fine-tuning the position (solution) based on the current position and the best fitness value of the particle; the social influence part, as indicated by the imitating behavior of the particles. Φi equals to c1*rand() and φg equals to c2*rand(), where c1 and c2 are accelerating constants; rand() stands for the random number generated from the uniform distribution between [0,1]. Equations (7), (8) and (9) update the position and velocity vectors.

Since the generated velocity vectors may be too large to control, a maximum velocity Vmax is defined to limit the range of the velocity vectors. In addition, the xij value should be kept between 0 and 1; therefore, the value of Vmax cannot be too large to reduce the solving power of the algorithm and the slower convergence.

4. Check the stop criterion

In order for the algorithm to better approximate the optimal value, the number of iterations plays a crucial role. Here the number of iterations is the stop criterion and more iteration is preferred; however, the efficiency may be compromised. Various iteration choices can be experimented to compare the convergence speed.

5. Find the optimal solution

The position (solution) calculated through PSO algorithm which satisfies both the stop criterion and constraints is converted to project direct cost and is regarded optimal.

3.2 Step 2: Calculate the completion probability through Monte Carlo simulation

Monte Carlo simulation is a widely used simple method based on the theory of large number. It is very suitable in finding the estimates of some non-linear, non-analytic problems. Although PSO algorithm can help in finding the optimal costs within specific time constraint, however, the optimal cost may hardly be achieved since it is pushing the envelope to the limit with little tolerance to the uncertainties. Thus Monte Carlo simulation is introduced in order to handle the uncertainties. Following are the applied steps using the Monte Carlo simulation.

For each time constraint:

- Set the total number of simulations N, the counter N' =0
- Set the time constraint
- Input the functions of time and cost for each project activity
- Generate the time and cost for each activity
- Calculate the total project duration
- If project duration ≦ time constraint, N' =N'+1, record the project direct costs
- Calculate Pc=N' / N, which is the completion probability for the specific time constraint

3.3 Step 3: Tabulate the probabilities for various time/cost combinations

A completion probability table is prepared from the values of the minimal project direct cost through PSO in step 1 and the ones with the simulation results from Monte Carlo simulation in step 2.

4. A Numerical Case

An underground sewage pipeline project constructed in Taiwan is used as a numerical case study. Figure 1 shows the project network.

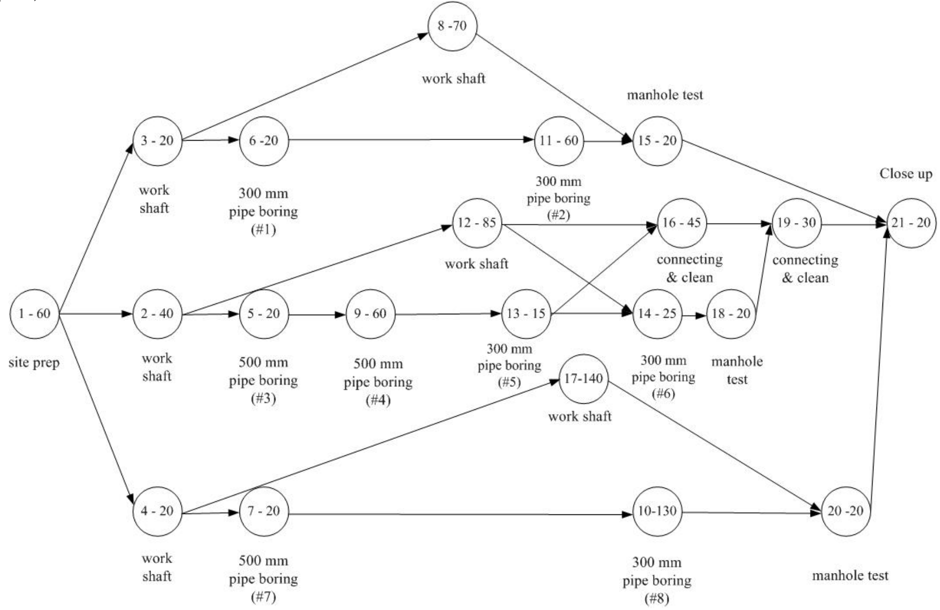


Figure 1. Network Diagram of the Underground Sewage Pipeline Project

Assume that the possible duration of this project are 220, 230, 240, 250, 260, 270, 280 and 290 days; as for the project costs, 3710, 3730, 3760, 3790, 3820, 3850, 3880, 3910, 3940, 3970 (in the unit of 10 thousand NT) are given. It is also common in project scheduling to assume the activity time distribution to be beta distribution. The time/cost relation is modeled as a quadratic function as proposed by Li et al. (1999). The quadratic functional form is that of $C(i)=aiT(i)^2+bi$, where C(i) is the cost and T(i) is the time; ai and bi are constants. The functional relationship for each activity has its own functional form is built up from field data. Finally, a complete table showing the relationship of every activity is established.

Step 1. Estimate the minimal project direct cost using PSO heuristic

Based on the assumed time durations of the project, PSO algorithm is performed with following settings: (iterations=500, number of particles=20, Vmax=5, initial inertia weight=1.5, final inertia=0.5, range of time duration: 240~290 days.) The minimal project direct costs for this time range are derived and the curve is drawn, as shown in Figure 2.

Step 2: Completion Probability by Monte Carlo Simulation

According to step 2 as mentioned in the previous section, number of simulations is set to be 10000, which can provide a precision of 4th digit.

Step 3：Tabulate the project completion probability for various time/cost combinations

As shown in Table 1, the probability matrix based on Monte Carlo simulation is generated. The value in the cell indicates the completion probability under certain time/cost constraint.

4.1 Case Analysis

The completion probability, as seen in Table 1, shows the chance of successful completion under certain time/cost constraint. Often, the project performer is more concerned about the profit margin, but overlooks the risk of its inability to carry out the project under tight time constraint. It is understandable that the more resources being put in the project, the faster the project can be carried out. However, as can be seen from Table, there is a region where hardly any time/cost combination can generate higher completion probability, or higher chance of success completion of project. This probability matrix can be further divided into three regimes: the first regime with cell values very close to zero and it indicates the situation that hardly these time/cost combinations should exist since the success chance is minimal; the second regime includes the cell value between (>0) and 0.6 and it indicates the success chance varies a lot from minimal to improving. Finally, the third regime includes cell values over 0.6 and it indicates very high success chance and promising to complete the project within these time/cost combinations. Therefore in the studied project, in order to have a successful and promising result, the time should be above 280 days and the costs should be above 39,100 thousand NT.

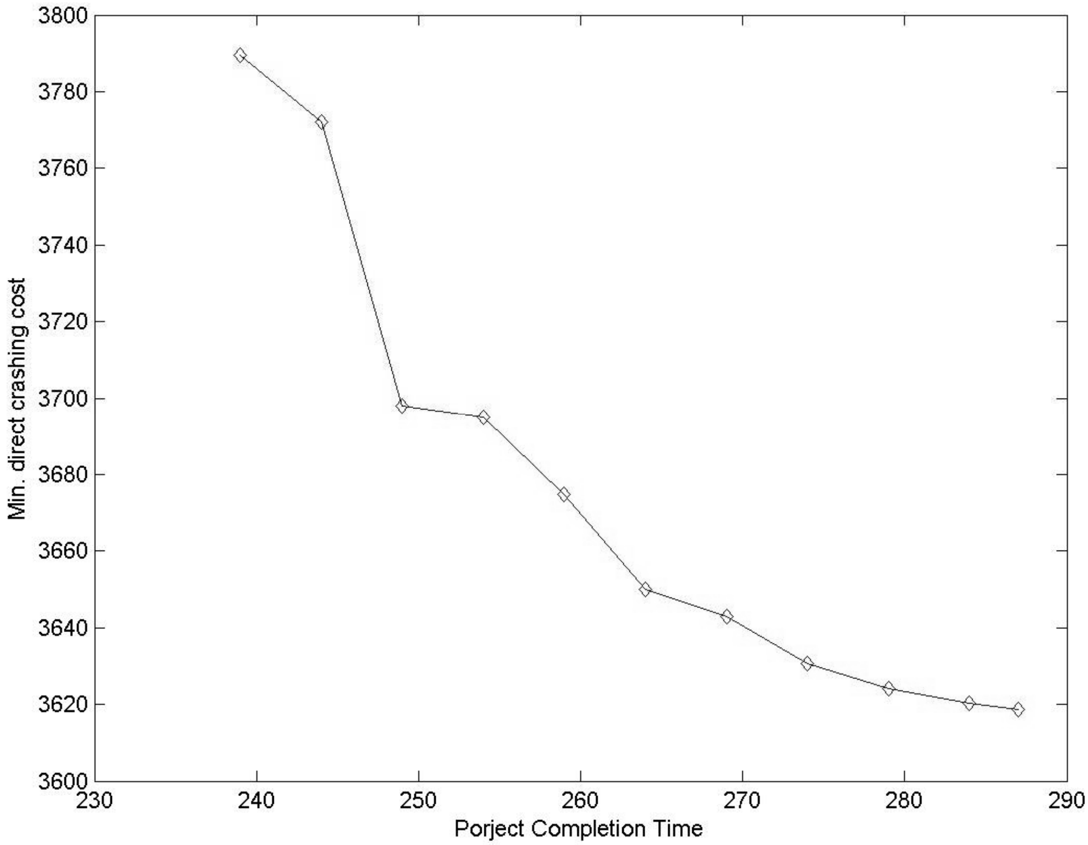


Figure 2. Activity Time and Cost Relationship (Beta Distribution)

Table 1. Completion Probability Matrix (Beta Distribution)

		Time Constraint (in days)							
		<=220	<=230	<=240	<=250	<=260	<=270	<=280	<=290
Project Direct Costs in 10,000 NT	<=3710	0	0	0	0	0	0	0	0
	<=3730	0	0	0	0	0.0002	0.0002	0.0002	0.0002
	<=3760	0	0	0	0	0.0025	0.0034	0.0034	0.0034
	<=3790	0	0	0	0.0014	0.0519	0.0669	0.0669	0.0669
	<=3820	0	0	0	0.0338	0.3425	0.3903	0.3904	0.3904
	<=3850	0	0	0	0.1557	0.7765	0.8398	0.8399	0.8399
	<=3880	0	0	0.0003	0.2348	0.9275	0.992	0.9921	0.9921
	<=3910	0	0	0.0005	0.2412	0.9354	0.9999	1	1
	<=3940	0	0	0.0005	0.2412	0.9354	0.9999	1	1
	<=3970	0	0	0.0005	0.2412	0.9354	0.9999	1	1

5. Conclusions and Recommendation

The study proposes a framework incorporating the PSO heuristic and Monte Carlo simulation to analyze the project completion probabilities among various time/cost constraints. From the numerical case study, it can be found that in the most cases, the project direct costs from PSO under time constraint are in the lower completion probability regime. Often the project direct costs derived from PSO are usually very optimistic in executing the projects.

In the study, it is assumed that the executing sequence of project activities is fixed and cannot be adjusted should there be resource constraints or fast-tracking. For project crashing problem, it is suggested that the resource conflicts and modified executing sequence can be further studied to fit into more realistic needs.

References

[1] Diaz, C. F., amf Hadipriono, F. C. (1993) "Nondeterministic networking methods" Journal of Construction Engineering and Management, 119(1), 40-57.

[2] Daisy X. M. Zheng, S. Thomas Ng, and Mohan M. Kumaraswamy. (2004) "Applying a genetic algorithm based multi-objective approach for time-cost optimization" Journal of Construction Engineering and Management, 130(2), 168-176.

[3] Ke, H. and Liu, B. (2005) "Project scheduling problem with stochastic activity duration times" Applied Mathematics and Computation 168, 342-353.

[4] Kelly, J. E., Jr. (1961) "Critical path planning and scheduling: mathematical basis" Operations Research. 9(3), 167-179..

[5] Laslo, Z. (2003) "Activity time-cost tradeoffs under time and cost chance constraints." Computers & Industrial Engineering 44, 365-384.

[6] Li, H., Cao, J. N., and Love, P. E. D. (1999) "Using machine learning and GA to solve time-cost trade-off problems" Journal of Construction Engineering and Management, 125(5), 347-353.

[7] Meyerm W. L., and Shaffer, L.R. (1963) "Extensions of the critical path method through the application of integer programming" Civil Engineering Construction Research Ser. 2, University of Illinois, Urbana.

[8] Patterson, James H. and Huber, Walter D. (1974). "A horizon-varying, zero-one approach to project scheduling" Management Science 20 (6), pp. 990-998.

[9] Yang, I T. (2006). "Performing complex project crashing analysis with aid of particle swarm optimization algorithm" International Journal of Project Management 25, 637-646.

Bridge Construction Progress Monitoring Using Image Analysis

Changyoon Kim[1], Hyoungkwan Kim[2] and Yeonjong Ju[3]

1School of Civil and Environmental Engineering, College of Engineering, Yonsei University, 134 Shinchon-dong, Seodaemun-gu, Seoul 120-749, Republic of Korea; PH (+822) 2123-7940 ; FAX (+822) 364-5300; e-mail: changyoonkim@yonsei.ac.kr
[2]Associate Professor, School of Civil and Environmental Engineering, College of Engineering, Yonsei University, 134 Shinchon-dong, Seodaemun-gu, Seoul 120-749, Republic of Korea; PH (+822) 2123-5799; FAX (+822) 364-5300; e-mail: hyoungkwan@yonsei.ac.kr
[3]School of Civil and Environmental Engineering, College of Engineering, Yonsei University, 134 Shinchon-dong, Seodaemun-gu, Seoul 120-749, Republic of Korea; PH (+822) 2123-7940 ; FAX (+822) 364-5300; e-mail: j970132@hotmail.com

Abstract

Construction progress monitoring is a critical task for construction managers. Traditionally, site superintendents walk the construction site to verify the progress report and understand the current state of construction processes. To get rid of the time consuming process, this research proposes a construction progress monitoring system comprised of a network type closed-circuit television (CCTV) camera, a wireless local area network, and image processing techniques. A civil engineering construction site was selected and the system is able to capture the images of the construction processes. In this paper, the hardware architecture of the system and its initial application results are presented with some promising outcomes.

Introduction

Construction progress monitoring is a critical task for construction managers. Successful progress monitoring prevents construction projects from falling into the pitfalls of ineffective management. Nowadays, many construction projects are provided with various photogrammetry equipments such as digital camera and closed-circuit television (CCTV) camera to record the pictures of construction sites. These saved images are typically used only for documentation of daily or weekly construction processes. Coupled with image processing techniques, the images can potentially be used to produce information of certain activity progress. In this paper, a research effort is introduced where a CCTV camera–based system is used to monitor the construction operations of a cable stayed bridge. The image processing system is developed in conjunction with the 3D (Three-Dimensional) CAD (Computer-Aided Design) model, in order to accurately and automatically track the construction progress level of the segment between two main pylons. To verify the construction progress monitoring methodology, tests are conducted at the cable-stayed bridge construction site.

Research Background

Image processing technique is gaining increased recognition in construction material engineering areas. Kim et al. (2002; 2003; 2004), with the use of laser technology, captured 3D image data of construction aggregates, and analyzed the images with advanced imaging algorithms to characterize the quality of the construction materials. Abdel-Qader et al. (2003) compared diverse image processing techniques chose a series of techniques that work best for crack identification in a bridge. By combination of image processing techniques and a database of construction materials, Brilakis and Soibelman (2008) used a shape retrieval mechanism to recognize a range of construction material resources.

On the other hand, image process techniques have been used for analyzing construction project processes. Zou and Kim (2007) suggested a strategy for analyzing idle time of hydraulic excavators, using the hue, saturation, and value color space. Quiñones-Rozo et al. (2008) applied image processing techniques for identifying and quantifying the excavating area. Bosche and Haas (2008) obtained promising results of progress monitoring using a 3D laser scanner and a simple 3D model. These efforts have shown a big potential for imaging techniques to produce the essential construction management information in an effective manner.

System Architecture for Wireless Image Data Acquisition

The images of the bridge construction processes are automatically transferred to the head office. Fig. 1 shows the system architecture for image data acquisition mechanism. A 704×480 network CCTV camera was installed at the cross beam of the pylon. This network CCTV camera can directly be connected to the internet, so that the acquired data are easily transferred to end users. However, due to the harsh construction environment, it is not easy to transfer the images through a direct LAN (Local Area Network) cable. It is possible that some heavy construction equipment or other construction vehicles damage the LAN cable.

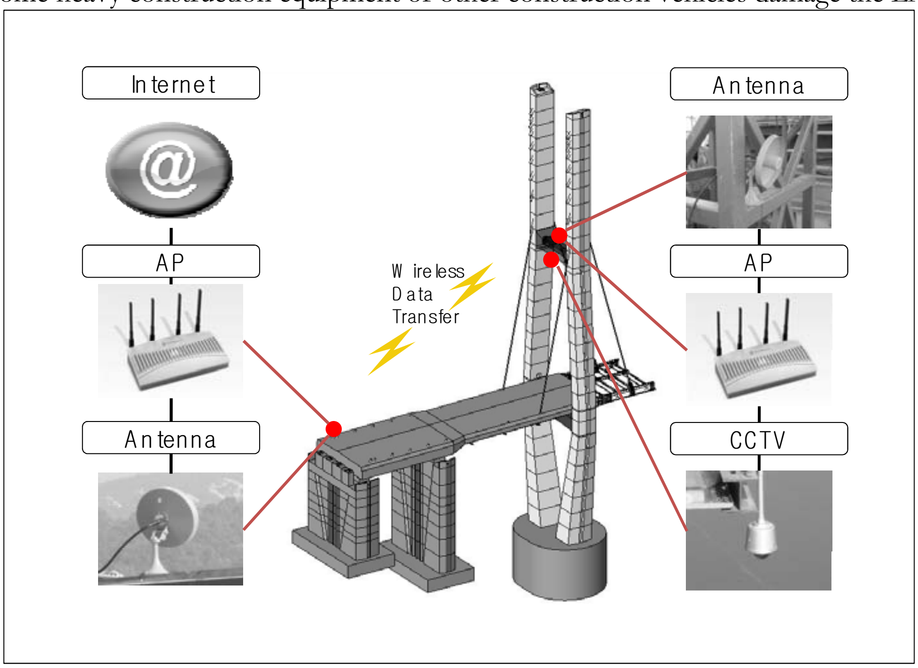
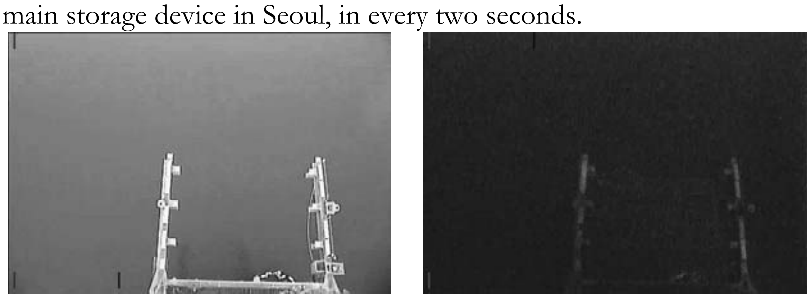


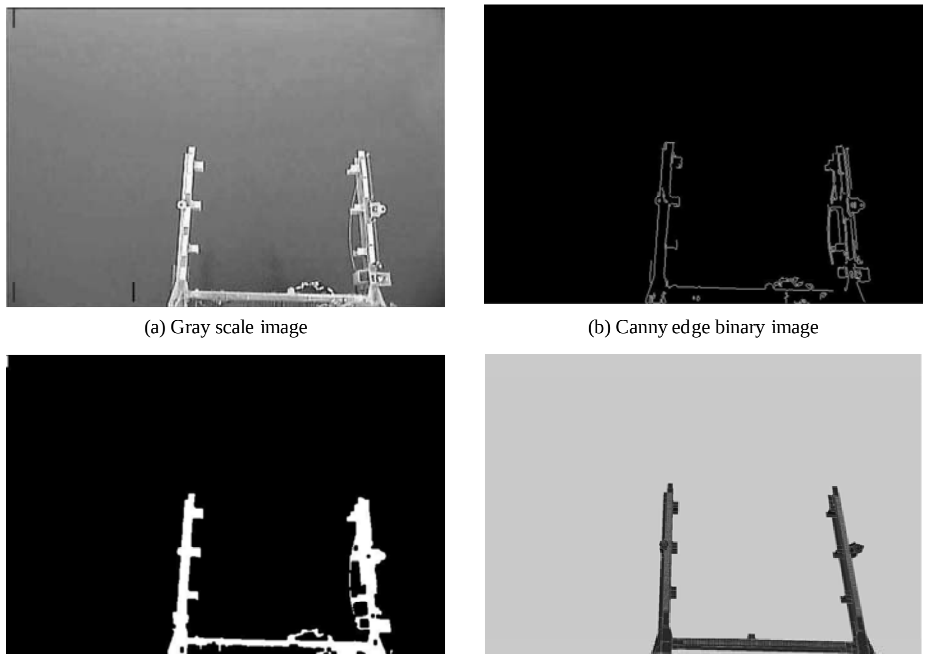
Fig. 1. Automated wireless image data acquisition system

Two access points (devices that allow for wireless communication) are used: one for data transmission and the other for data reception. One access point was installed to be connected to the camera at the cross beam of the pylon and the other to be connected to the internet at the entrance of the construction site (the end of the side span), resulting in the establishment of the Wireless LAN (WLAN) environment. To cover the communication distance of about 68.4 m between two access points, two directional antennas were used. The image data which acquired by the CCTV camera was successfully transferred through the WLAN system. Construction managers at the head office in a remote area can monitor the construction progress of the cable stayed bridge in a real time via the internet connection. For documentation purpose, the images are being stored at the main storage device in Seoul, in every two seconds.



(a) Day (b) Night

Fig. 2. The view of construction site: (a) Day and (b) Night



(a) Gray scale image

(b) Canny edge binary image

(c) Canny edge binary image with holles filled

(d) 3D model

Fig. 3. The images of processed data and 3D model

Object Segmentation

A total of 14 structural components comprise one segment of the main span: two plate girders, three floor beams, three stringers, and six concrete panels. These 14 structural components are sequentially moved and installed by a derrick crane. Fig. 2 shows the day and night view of the process of two plate girders installation.

The acquired image data of the two plate girders were processed and analysed on a Matlab™ platform. Fig. 3 presents the processed images and the 3D CAD model of structural components. To perform the edge detection of the object of interest, the original RGB (Red, Green, and Blue) color image (Fig. 2 (a)) was converted to a gray scale image. The resultant image of the conversion process is presented at Fig. 3 (a). Next, the edge of the object of interest was detected by the canny edge detection algorithm using a 5×5 Gaussian filter with automatically selected optimum high threshold value T_H and low threshold value T_L. The canny edge binary image of the two plate girder is shown in Fig. 3 (b). The gaps and holes of the detected edge were closed and filled with binary morphological closing and filling functions. The result of these two functions is presented at Fig. 3 (c). To compare the actual progress to the original plan, the 3D CAD model of the cable stayed bridge was generated in the Autodesk Revit Architecture Suite environment. By matching the processed image data (Fig. 3 (c)) with the 3D model (Fig. 3 (d)), the progress of the construction process can be measured.

Conclusions

This research presents the whole process of monitoring a cable stayed bridge construction, from the automated data acquisition to the data analysis. With the use of a CCTV camera and WLAN, the image data of the construction process were automatically transferred from the construction site to the head office storage device. To analyse the construction progress, image processing techniques along with the 3D CAD model were used for the comparison between the acquired data and as-planned data.

This preliminary research shows the possibility of automated process monitoring for civil structures such as cable stayed bridges using CCTV cameras and image processing techniques. Future study will focus on how to extract useful management information from the processed image data. This methodology is expected to lead to a fully automatic bridge construction progress monitoring system for effective construction management.

Acknowledgement

This work was supported by a Grant (KRF-2008-331-D00611) from the Korea Research Foundation and a Grant (07UrbanRenessanceB03) from High-Tech Urban Development Program funded by the Korean Ministry of Land, Transport, and Maritime Affairs.

References

[1] Abdel-Qader, I, Abudayyeh, Osama, and Kelly, M. E. (2003). "Analysis of Edge-Detection Techniques for Crack Identification in Bridge." Journal of Computing in Civil Engineering, 17(4), 255-263.

[2] Bosche, F., and Haas, C. T. (2008). "Automated retrieval of 3D CAD model objects in construction range images." Automation in Construction, 17(4), 499-512.

[3] Brilakis, I. K., and Soibelman, l. (2008). "Shape-Based Retrieval of Construction Site Photographs." Journal of Computing in Civil Engineering, 22(1), 14-20.

[4] Kim, H., Haas, C. T., Rauch, A. F., and Browne, C. (2002). "Dimensional Ratios for Stone Aggregates from 3D Laser Scans." Journal of Computing in Civil Engineering, 16(3), 175-183.

[5] Kim, H., Haas, C. T., Rauch, A. F., and Browne, C. (2003). "3D Image Segmentation of Aggregates from Laser Profiling." Computer-Aided Civil and Infrastructure Engineering, 18(4), 254-263.

[6] Kim, H., Haas, C. T., and Rauch, A. F. (2004). "Automated Quality Assessment of Stone Aggregates Based on Laser Imaging and a Neural Network." Journal of Computing in Civil Engineering, 18(1), 58-64.

[7] Quiñones-Rozo, C. A., Hashash, Y. M. A., and Liu L. Y. (2008). "Digital Image Reasoning for Tracking Excavation Activities." Automation in Construction, 17(5), 608-622.

[8] Zou, J., and Kim, H. (2007). "Using HSV Color Space for Construction Equipment Idle Time Analysis." Journal of Computing in Civil Engineering, 21(4), 238-246.

Resource Management in Civil Construction Using RFID Technologies

Changyoon Kim[1], Yeonjong Ju[2], Hyoungkwan Kim[3] and Jung-Hoon Kim[4]

[1]School of Civil and Environmental Engineering, College of Engineering, Yonsei University, 134 Shinchon-dong, Seodaemun-gu, Seoul 120-749, Republic of Korea ; PH (+822) 2123-7940 ; FAX (+822) 364-5300 ; e-mail: changyoonkim@yonsei.ac.kr
[2]School of Civil and Environmental Engineering, College of Engineering, Yonsei University, 134 Shinchon-dong, Seodaemun-gu, Seoul 120-749, Republic of Korea ; PH (+822) 2123-7940 ; FAX (+822) 364-5300 ; e-mail: j970132@hotmail.com
[3]Assistant Professor, School of Civil and Environmental Engineering, College of Engineering, Yonsei University, 134 Shinchon-dong, Seodaemun-gu, Seoul 120-749, Republic of Korea ; PH (+822) 2123-5799 ; FAX (+822) 364-5300 ; e-mail: hyoungkwan@yonsei.ac.kr
[4]Assistant Professor, School of Civil and Environmental Engineering, College of Engineering, Yonsei University, 134 Shinchon-dong, Seodaemun-gu, Seoul 120-749, Republic of Korea ; PH (+822) 2123-5804 ; FAX (+822) 364-5300 ; e-mail: junghoon@yonsei.ac.kr

Abstract

Large amount of construction resources and their scattered locations in a civil construction site make it extremely difficult for project managers to effectively utilize the resources. In order to manage construction resource effectively, a Radio Frequency Identification (RFID) technology was used in this research. RFID tags for metallic objects were attached to structural components used in a cable-stayed bridge. The RFID reader and antennas were installed at the construction site for tracking the structural components and the acquired data were transferred to the head office via Local Area Network (LAN). The proposed RFID system enabled real-time monitoring of the construction materials in the civil construction site.

Introduction

Large amount of construction resources and their scattered locations in a civil construction site make it extremely difficult for project managers to effectively utilize the resources. For this reason, there are strong needs for managing construction resources in a real-time manner (Kim et al., 2008). Diverse sensing techniques such as Radio Frequency Identification (RFID) technology and flash Laser Radar (LADAR) were being used in construction sites (Song et al. 2006; Kim et al. 2008). The RFID technology is a wireless object identification method relying on radio frequency signals. A RFID reader transmits the signal through the Transmission (TX) antenna and receives it which is reflected by the RFID tag through the Reception (RX) antenna. In this research, a preliminary experiment was conducted to characterize a RFID system composed of readers and tags. Capability of the reader and tags for metal and non-metal objects were identified using the frequency of 910.4 ~913.6Mhz UHF (Ultra High Frequency), and off the shelf tags classified as Electronic Product Code (EPC) class 1, Generation 2 (GEN Ⅱ) RFID. Finally, a case study was conducted to confirm the applicability and robustness of the RFID system in a real civil construction site.

Research Background

Pioneering researchers have tried to use RFID technologies for the contactless identification of construction resources in various construction sites, including plant and building construction. However, RFID application in civil engineering works such as bridge construction has not yet been reported. Song et al. (2006) suggested a strategy for tracking locations of construction materials using the RFID technology. In their study a construction vehicle was first located, and By using Global Positioning System (GPS), RFID, and Bluetooth communication protocol (IEEE 802.15.1), construction vehicles were automatically tracked. The RFID technology is often used for facility management and maintenance. Ergen et al. (2007) presented a RFID application strategy for management of fire valves. In order to record inspection history of fire valves, RFID tags were used. Ko (2008) developed a web-based facility management information system, along with the RFID technology, to manage and maintain building facility. This is a trend found in other research efforts, too. Through the combination of RFID and 4D CAD (Computer Aided Design), construction progresses can automatically be monitored and visualized (Chin et al., 2008).

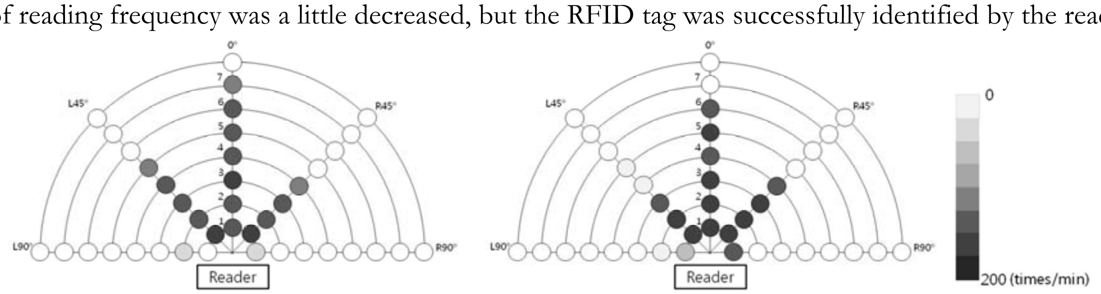
Preliminary Experiment

The drawback of the normal RFID system is that the tags are not easily detected in a metallic environment (Kim et al. 2008). The radio frequency signal is interfered by metallic materials so that RFID tags cannot reflect the signal that transmitted from the TX antenna. This brings about serious problem because the cable stayed bridge, chosen as a test case for RFID applicability, consists of many metallic structural objects.

To overcome this shortcoming, metal tags, which were specially designed for detecting metal objects, were used. The metal tag is composed of RFID chip, antenna, and an insulator. Insulators create an extra layer between the RFID tag (chip and antenna) and the object that construction managers want to identify. This is to satisfy the condition where the radio frequency signal is not interfered by the metallic environment.

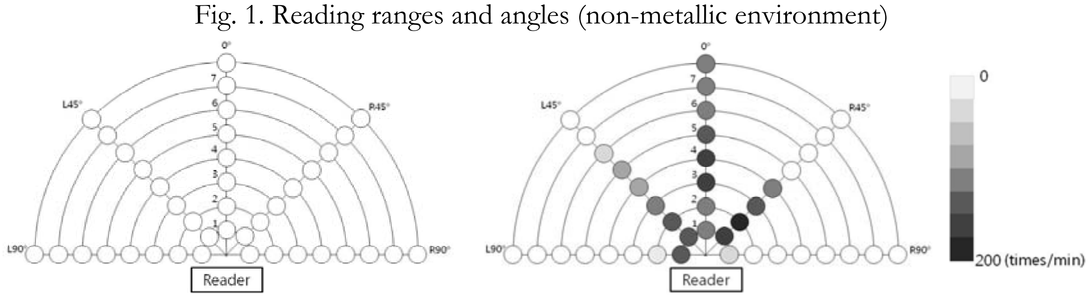
To measure the readability of RFID tags, the number of tag detection during a specified time duration (1 min) was counted. In addition, the reading ranges (distance) and angles were also measured. The experiment of reading ranges and angles was repeated three times and the average value was used for the analysis. Fig. 1(a) shows the experiment results for non-metal tags. When the traditional RFID tag was placed in front of the reader (with the angle of 0°) in a non-metallic environment, the maximum reading range was seven meters from the RFID reader. When the traditional RFID tag was placed on the oblique line (with the angles of Left (L) 45° and Right (R) 45°), the maximum reading range was four meters and five meters, respectively. The reading ranges of the tag were only one meter and two meters at the angles of R90° and L90°, respectively.

A metal RFID tag was also used in this experiment. The reading range of the tag was almost same as the traditional RFID tag when it comes to non-metallic objects. However, the result was quite different for metallic objects. Fig. 2 shows the experimental results of reading ranges and angles of the metal tag. None was identified when the traditional RFID tag was attached to the metallic materials. However, the RFID tag for metallic environment was well identified even when the tag was attached to the metallic materials. The level of reading frequency was a little decreased, but the RFID tag was successfully identified by the reader.



(a) Non-metal Tag (a) Metal Tag

Fig. 1. Reading ranges and angles (non-metallic environment)



(a) Non-metal Tag (a) Metal Tag

Fig. 2. Reading ranges and angles (metallic environment)

Case Study and System Architecture

A case study is being conducted in a cable-stayed bridge construction site in Chungcheong-bukdo, Jecheon, Korea. Precast concrete panels that will comprise the upper deck of the bridge segment was chosen as the material type to be monitored (Fig. 3). The panels are being fabricated in a lot approximately 50 m away from the main entrance of the construction site. Since the panels consists of cement, aggregates, and reinforcing bars which are metal, the metal RFID tags are used for successful identification. RFID tags that were purchased were first tested in an indoor environment before being deployed in the site. 19 RFID tags out of 197 (9.64%) turned out to be defective, and the defective tags are not used for the case study. Since the concrete panels are still being fabricated, not all the non-defective tags are deployed yet. The RFID technology-based material tracking system is now in place and the initial tracking result is coming out.

To facilitate successful material management, Radio Frequency Identification (RFID) and Local Area Network (LAN) are used in this study (Fig. 4). To identify construction materials with RFID tags at the construction site, two pairs of antennas (two TXs and two RXs) and a RFID reader were installed at the entrance of the construction site. Two pairs of antennas (TX and RX antenna) are installed to broaden the reading range of RFID readers. When the material with an RFID tag enters into the detection zone of the reader, it can be identified and the information stored in the tags can be transferred to the main database server via LAN. A preliminary database was developed using MS-Access and the VDSL (very high-data rate digital subscriber line) is used for the internet connection.

Fig. 3. Precast concrete panels

Conclusions

This research presents an effort to evaluate the applicability of the RFID technology in managing resources in civil construction sites. Coupled with the LAN connection, the system is able to provide the management with real time material tracking information in any place where the internet connection is available, including construction site office and head office. The special tags for a metallic environment were also adopted to enhance the robustness of the RFID systems in the cable-stayed bridge construction site where diverse construction materials were used.

This study is unique and contributory to the existing body of knowledge in the sense that to the knowledge of the authors, this is the first trial where RFID technology is used in a civil construction. Although RFID technologies have shown big potentials in architectural engineering or plant constructions, no significant effort has yet been reported with regard to the RFID application in civil construction sites. Future study will focus on extracting statistically meaningful and reliable conclusions as to how best to be able to utilize RFID technologies in civil engineering construction such as a cable stayed bridge.

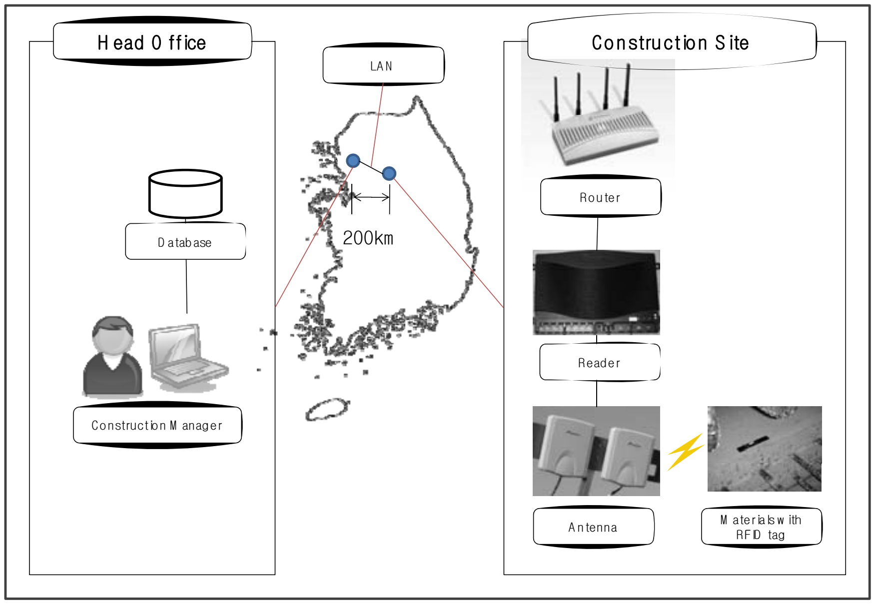


Fig. 4. Resource management system architecture

Acknowledgement

This work was supported by a Grant (KRF-2008-331-D00611) from the Korea Research Foundation and a Grant (07UrbanRenessanceB03) from High-Tech Urban Development Program funded by the Korean Ministry of Land, Transport, and Maritime Affairs.

References

[1] Chin, S., Yoon, S., Choi, C., and Cho, C. (2008). "RFID + 4D CAD for Progress Management of Structural Steel Works in High-Rise Buildings." Journal of Computing in Civil Engineering, 22(2), 74-89.
[2] Eegen, E., Akinci, B., East, B., and Kirby, J. (2007). "Tracking Components and Maintenance History within a Facility Utilizing Radio Frequency Identification Technology." Journal of Computing in Civil Engineering, 21(1), pp. 11-20.
[3] Kim, C., Kim, H., Ryu, J., Kim, C., Kim, M. K. (2008). "Ubiquitous Sensor Network for Real-time Construction Material Monitoring." Proceedings of International Conference on Computing in Civil and Building Engineering (ICCCBE 2008), Tsinghua University, Beijing, China.
[4] Kim, C., Kim, H., Han, S. H., Kim, C., Kim, M. K., and Park, S. H. (2008). "Developing a Technology Roadmap for Construction R&D through Interdisciplinary Research Efforts." Automation in Construction, In press.
[5] Kim, C., Son, H., Kim, H., and Han, S. H. (2008). "Applicability of Flash LADAR to 3D Spatial Information Acquisition and Modeling on a Construction Site." Canadian Journal of Civil Engineering, 35, 1131-1341.
[6] Ko, C.-H. (2008). "RFID-based Building Maintenance System." Automation in Construction, In press.
[7] Lu, M., Chen, W., Shen, X., and Lam, H.-C., and Liu, J. (2007). "Positioning and Tracking Construction Vehicles in Highly Dense Urban Areas and Building Construction Sites." Automation in Construction, 16(5), 647-656.
[8] Song, J., Haas, C. T., and Caldas, C. H. (2006). "Tracking the Location of Materials on Construction Job Sites." Journal of Construction Engineering and Management, 132(9), 911-918.

RFID-Assisted Lifecycle Management of Building Components Using BIM Data

Ali Motamedi[1] and Amin Hammad[2]

[1]Research Assistant, Concordia Institute for Information Systems Engineering,
Concordia University, 1455 De Maisonneuve, Montreal, QC, Canada H3G 1M8; PH (514) 848-2424, ext. 7041; FAX (514) 848-3171; e-mail: a_mot@ciise.concordia.ca
[2]Associate Professor, Concordia Institute for Information Systems Engineering,
Concordia University, 1455 De Maisonneuve, Montreal, QC, Canada H3G 1M8; PH (514) 848-2424, ext. 5800; FAX (514) 848-3171; e-mail:hammad@ciise.concordia.ca

Abstract

The AECOO industry is highly fragmented; therefore, efficient information sharing and exchange between various players are evidently needed. Furthermore, the information about facility components should be managed throughout the lifecycle and be easily accessible for all players in the AECOO industry. BIM is emerging as a method of creating, sharing, exchanging and managing the information throughout the lifecycle between all the stakeholders. RFID, on the other hand, has emerged as an automatic data collection and information storage technology, and has been used in different applications in AECOO. This research proposes permanently attaching RFID tags to facility components where the memory of the tags is populated with accumulated lifecycle information of the components taken from a standard BIM database. This information is used to enhance different processes throughout the lifecycle. A conceptual RFID-based system structure and data storage/retrieval design are elaborated. To explore the technical feasibility of the proposed approach, two case studies have been implemented and tested.

Keywords: RFID, Lifecycle Management, BIM, Construction Automation

Introduction

Radio frequency identification (RFID) is a type of automatic identification technology in which radio frequencies are used to capture and transmit data. It acts as electronic labelling and data-collection system to identify and track items. RFID based systems have been used in different applications in construction and maintenance, such as component tracking and locating, inventory management, equipment monitoring, progress management, facilities and maintenance management, tool tracking and quality control (Motamedi and Hammad, 2009). However, each of the above-mentioned applications is designed for only one specific stage of the facility lifecycle to serve the needs of only one of the stakeholders in a fragmented fashion, i.e., Architects, Engineers, Constructors, Owners and Operators (AECOO). This would increase the cost and the labor for adding and removing different tags at different stages and eliminate the chance of using shared resources among the stakeholders causing duplication of efforts and resources.

This research proposes permanently attaching tags to components in the manufacturing stage as an integrated part of the components. Having the tags permanently attached, where the information on the tags is gradually updated with accumulated lifecycle information, is beneficial for all the stakeholders throughout the stages of the lifecycle, from procurement and supply chain management to maintenance and disposal.

The use of attached RFID tags for lifecycle management has been proposed in the aerospace industry for storing unique ID and important lifecycle information on tags attached to aircraft parts for enhancing inspection and repair processes (Harrison et al., 2006). Ergen et al. (2007) proposed using RFID tags attached to engineered-to-order (ETO) components and explored the technical feasibility of such system by analysing component-related information flow patterns in ETO supply chains. They noted that integration of the data accessed with the broader information systems used across diverse organizations is an issue that needs to be investigated.

This paper aims to propose techniques to manage components' lifecycle data as well as extending the idea of attaching RFID tags to other types of engineered components within a constructed facility (i.e., made-to-order and off-the-shelf components). We also propose to include broader data types on the RFID tags that are attached to building components and are spread in a building. In the proposed approach, the information on the tags represents chunks of the Building Information Model (BIM) as a distributed database. This coupling between the BIM and the RFID information would allow reconstructing the database of the BIM (or part of it) based on the pieces of information distributed in all the attached tags. The proposed approach is further explored in two case studies of a high-rise building by deploying RFID tags on selected components for improving supply chain management, locating items, installation and maintenance activities, as well as progress management and visualization.

Review of Related Research

Radio Frequency Identification

RFID tag is a memory storage device for storing a certain amount of data. This information can be read wirelessly providing the ability to process large volumes of multiple data sets simultaneously. A basic RFID system consists of three components: an antenna, a transceiver (with decoder) and a transponder (RF tag) electronically programmed with information. The antenna can be packaged with the transceiver and decoder in order to become a reader. The reader can be configured either as a handheld or a fixed-mount device. RFID tags differ in many aspects, such as power source, frequency, readability range data transfer rates, data storage capacity, memory type, size, operational life, and cost (aimglobal.com, 2008).

While RFID technology has significant beneficial applications in manufacturing, retailing, transport and logistics industries, its potential applications in the AECOO industry have only begun to be explored (Song et al., 2006). The main usage of RFID is in supply chain and management and the tracking of materials, components, workers and equipment in construction projects. However, some researchers have proposed using RFID for progress monitoring, visualisation, quality control, and tracking components during inspection and maintenance activities (Motamedi and Hammad, 2009).

Building Information Model

The AECOO industry is highly fragmented in nature. This situation has resulted in significant barriers to communication between the various stakeholders, which in turn has significantly affected the efficiency and performance of the industry. Consequently, there is an evident need for a standard information transfer model between different software applications used in the AECOO industry. The BIM has been developed in order to tackle the problems related to interoperability and information integration by providing effective management, sharing and exchange of a building information through its entire lifecycle. BIM is extensible, open and vendor neutral and BIM data can be stored as a digital file or in a database, and can be shared and exchanged between several applications (Isikdag et al., 2007).

The Industry Foundation Classes (IFC) standard has matured as a standard BIM. IFC is an object-based, non-proprietary building data model and data exchange format. Completion of the IFC model facilitated the development of exchange standards. The Facility Information Council of the National Institute of Building Sciences (NIBS) formed NBIMS group aiming to speed the adoption of an open-standard BIM through the definition of information exchange standards based on the IFC model (East and Brodt, 2007).

Construction industry contracts require the handover of various documents. IFC-mBomb project demonstrated an approach for data capturing during design and construction, and data handover to facility operators (Stephens et al., 2005). Construction Operations Building Information Exchange (COBIE) project was initiated under NBIMS support with the objective of identifying the information exchange needs of facility managers and operators of data available upstream in the facility lifecycle (East and Brodt, 2007). While COBIE is designed to work with the BIM, COBIE data may also be created and exchanged using simple spreadsheets.

Proposed Approach

The lifecycle of a building can be divided into different stages where each stage is generally managed independently while exchanging partial information with other stages. The information related to each component should be tracked separately throughout the lifecycle. Furthermore, the information should be in a convenient format and stored at a suitable location to enable all the stakeholders to efficiently access throughout the lifecycle. Centrally stored information that is accessible electronically over a computer network is a solution for data access. However, having real-time access to information could be difficult since reliable connections to the central data storage may not be always available.

This research proposes adding structured information taken from BIM database to RFID tags attached to the components. Having the essential data related to the components readily available on the tags provides easy access for whoever needs to access the data regardless of having real-time connection to the central database or having a local copy of the required information.

System interaction design

In our proposed approach, every component is a potential target for tagging. Having standard tags attached to components would result in a massive tag cloud in the building. While having tags attached to all components would not happen in the immediate future, in order to benefit from the concept of having identity and memory tags on a mass of items, the subset of components to be tagged can be selected based on the scale of the project, types and values of the components, specific processes applied to these components, and the level of automation and management required by the facility owners.

The system design, including the data structure model and data acquisition method, is general for all components. The target components are tagged during or just after manufacturing and are scanned at several points in time. The scan attempts are both for reading the stored data, or modifying the data based on the system requirements and the stage at which the scan is happening. The scanned data are transferred to different software applications and processed to manage the activities related to the components. Fig. 1 shows the conceptual design for interaction between different system components. The generic software application communicates with RFID tags by using the reader API and stores and retrieves the lifecycle data to/from a central BIM database. The RFID specific information is added to the BIM database in the design phase as part of the product information.

The memory of the tag contains a subset of BIM information. While the BIM database is being populated by information by different software applications throughout the lifecycle, the tag memory space is modified and updated as the component is scanned. Fig. 2 conceptually shows how BIM data chunks are stored on tags attached to the building components. While the information is centrally stored in the BIM database, software applications copy the necessary information from the database to the memory space on the tags.

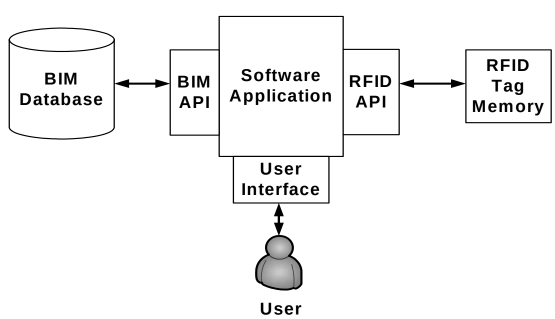


Figure 1 Conceptual system interaction design

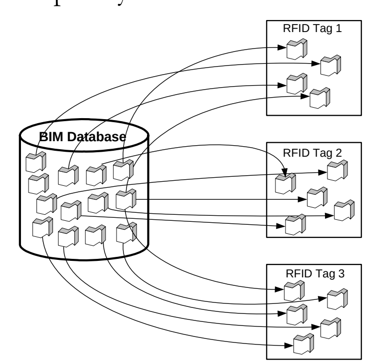


Figure 2 Conceptual BIM-Tag data relationship

Data capture methods

The structured data stored on the tags should be read, updated and changed by several RFID-based systems during the lifecycle. These modifications are executed by different types of RFID readers (stationary or mobile). In order to identify the suitable type of reader for each scan attempts, the detailed process requirements should be captured, such as the readability range, data transfer rate and portability. Moreover, the reader should be selected considering the type of the component (i.e., fixed, movable and temporary).

The data stored on the components can be read from different distances. The maximum readability distance depends on various factors, such as power level of reader, antenna type and size, frequency range and environmental factors. In some applications, it is desirable that the data be read from far distance. Hence, the system can detect the component even if the component is hidden or not visible. Other applications may require shorter readability. For example, if the tags are used to facilitate inspection activities, having short read/write range would guaranty that the inspector was in the required proximity of the component.

In the proposed approach, RFID tags are fixed to components; therefore, tags should be designed to have the maximum possible range and protection from noise and interference. However, it is always possible to control read/write range of the reader based on the process requirements.

Conceptual data structure

Considering the limited memory of the tags, the subset of BIM data stored on the tags has to be chosen based on the requirements. While data on a tag are changing during the lifecycle of the component and different software applications use and modify the data with different designated access levels, the memory of the tag should be virtually partitioned in a structured fashion based on predefined data types. `We propose to virtually partition the memory space into the following fields:

ID: In order to look up the component in the BIM database, there is a need to have a none-changeable, unique identifier (ID) for each component (e.g., EPCglobal (Electronic Product Code) Tag Data Standards).

Specifications: This field is dedicated to specifications of the component derived from the design and manufacturing stage of the lifecycle. Safety related information and hazardous material information are examples of *specifications*.

Status: Status field identifies the current main stage (e.g., in service, installed, manufactured, and assembled) and sub-stage (e.g., in service: waiting for inspection) of lifecycle of the component. The *status* information is used to decide which software application can use and modify the data in the *process data* field.

Process data: This field is relatively large compared to the other fields and is designed to store the information related to the component's current stage of the lifecycle. The data related to current processes to be stored on the tags are different and should be changed during the lifecycle. For example, assembling instructions are used only in the assembly stage. Therefore, the *process data* field contains only information related to the current lifecycle stage taken from BIM database. Moreover, the ownership (ability to read, modify or change) of the *process data*, should be restricted to one or a group of applications (e.g., inspection management software, installation management software) that are involved in that specific stage. The ownership of the *process data* field is decided based on the s*tatus* field as explained above. Fig. 3 shows how different software applications modify the *process data* field. Different applications use the same memory space but at different lifecycle stage. Fig. 3 demonstrates a sample component that follows a specific lifecycle pattern where BIM information is copied by different software applications on the memory of its RFID tag.

History data: This field is designated for storing the history data used during the lifecycle for maintenance and repair purposes. The history records are derived from BIM and accumulated during the lifecycle to be used in forthcoming stages.

Environment data: This field is designated for storing environment specific data, such as the location or the functionality and specifications of the space (e.g., floor plan). *Environment data* is also taken from BIM and contains all the information that is not related to the component itself.

BIM-Tag data exchange method

The proposed approach suggests using the attached RFID tags as a media for storing the structured information and providing a distributed data storage of BIM information. According to the fact that the media for storing the data is transparent to the efforts to collect, manage and share data, the data can be stored in a central database, RIFD tags, or in printed documents based on a data management method.

Several research projects are undergoing on the subjects of lifecycle information management and identification of information exchange paths, identification of crucial data that should be available during the lifecycle and handover methods between AECOO participants, the results of these researches can be incorporated in our proposed approach.

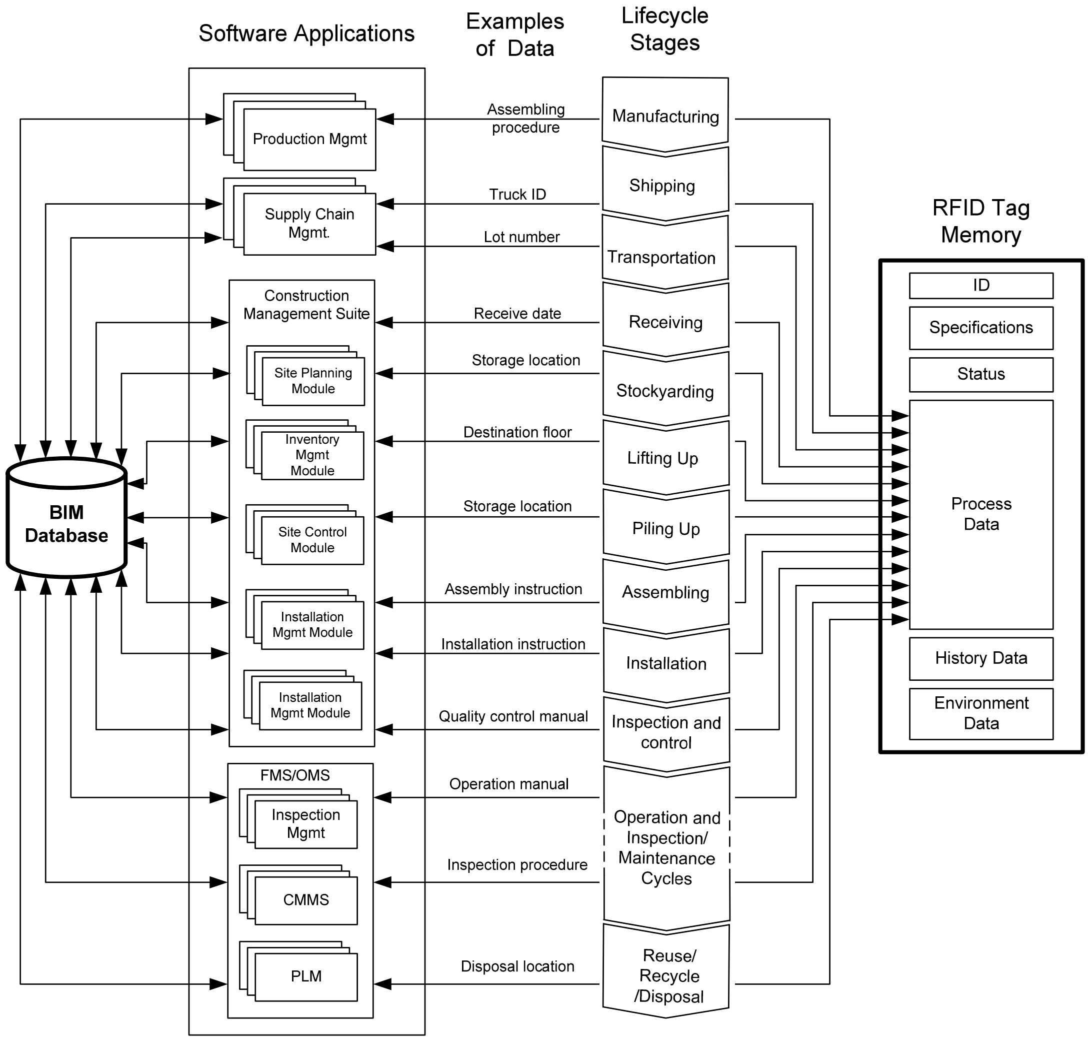


Figure 3 Process data update

Challenges

Although the proposed approach can be implemented using available hardware, due to high implementation and customization costs, it is not financially feasible at present. The challenges can be categorized under the following main topics:

Challenges related to adopting RFID technology: *(1) RF challenges* related to properties of magnetic waves and the effects of materials on them; (2) Lack of complete and international *standards*; (3) *Cost* of tags and infrastructure implementation; (4) *Data security* and privacy; (5) *Ruggedness* of tags that can operate in harsh environments for the construction industry; (6) *Data transfer speed*; (7) *Interoperability:* standards to cover all types of tags and frequencies and multi-protocol tags and readers; (8) *Power:* low power RFID systems to extend the lifetime; and (9) *Environment:* tags made of new materials to facilitate recycling.

Challenges in extending BIM and its implementation: The efforts for developing BIM standards are in their early stage and the available standards are not complete and thorough. Moreover, adopting BIM standards has its own challenges and obstacles; issues such as industry acceptance, change management from conventional methods to new BIM, qualified human resources, legal considerations and initial cost to change (hardware, software, training and implementation) have to be tackled for industry-wide implementation of BIM.

Technology adoption and social challenges: Wide Implementation of such systems would bring resistance from companies that are using traditional methods because of needed extra efforts and training.

Process related challenges: The processes involved in building lifecycle should be reviewed and re-engineered considering new opportunities.

Case Study 1: Progress management and 4D visualization

This case study is designed to facilitate the process of progress monitoring of construction projects and to provide visualisation aid for component status tracking. The result of implementing the case study is accurate progress measurement data resulting in accurate 4D model and 3D visualization based on component's status.

The prototype system is composed of six subsystems: (1) the database that store the data extracted from the BIM, which will be updated by RFID reads and other software updates (e.g., inspection data), (2) the 3D modeling software that stores the data in IFCxml format, (3) the scheduling software, (4) the 4D simulation software, (5) the FM software, and (6) the RFID reader interfaces. The communications between the subsystems are based on standard protocols providing scalability and interoperability. The 4D simulation software obtains the geometrical information from the 3D software and the timing and status information from the database to produce different real-time views of the facility using a predefined colouring scheme. These views help project managers and the FM team to better visualize the status of the facility.

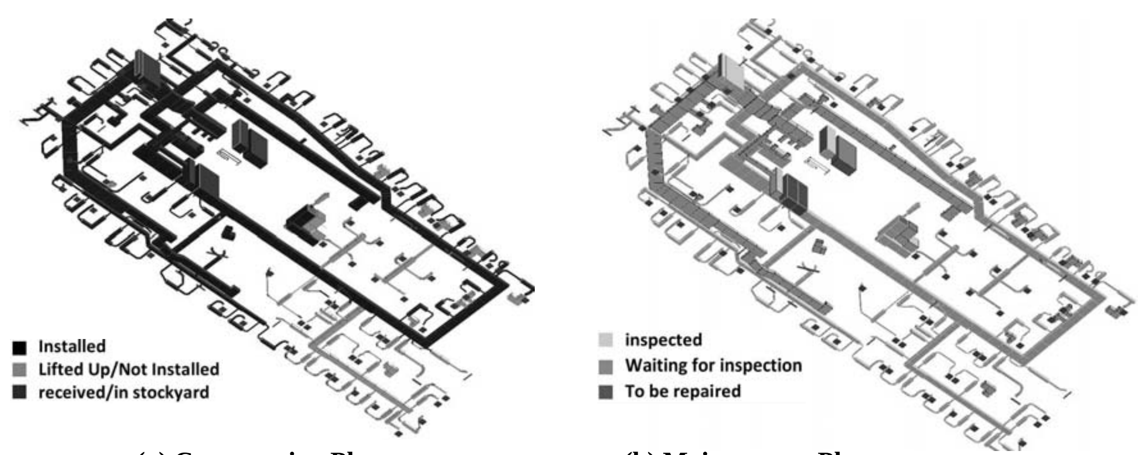
In this study, we focused on the progress monitoring and lifecycle management of the HVAC components in the new building of John Molson School of Business (JMSB) at Concordia University. Various active RFID tags are used that operate in UHF frequency and have 8 or 32 KB of memory. The tags are updated by a mobile reader attached to the inspector's or maintenance worker's hand-held device.

Fig. 4 shows sample snapshots of 4D visualization of the HVAC system on the 14^{th} floor of the building. Fig. 4(a) shows the status of the components during the construction phase. The components that are installed are shown in black. The components that are lifted up but not yet installed are shown in red. The components that are in the stockyard and have not been lifted up are shown in grey. Fig. 4(b) shows the status visualization during the operation phase. The component in green is waiting for inspection and the component in dark brown needs to be repaired.

Case Study 2: Fire equipment inspection and maintenance

In this case study, RFID tags are used for storing information about fire safety equipments. Amongst all safety related equipments, fire extinguishers and safety valves are chosen because of their importance and the higher frequency of their maintenance activities.

In our prototype system, crucial information is stored on tags attached to the extinguishers and valves. This would provide the information about the history and the condition of the extinguishers and valves for inspectors and maintenance/repair personnel without access to any central database.



(a) Construction Phase **(b) Maintenance Phase**

Figure 4 HVAC 3D drawings of one floor of the JMSB building

Two different types of tags have been tested and used in the prototype system: Active tags with 8 or 32 KB of memory and standard passive tags with 96 bits of memory. The active tags are long range but the passive tags have the readability range of few inches for a typical handheld reader.

Short write distance for tags would guaranty that the inspector did the inspection and maintenance activity in close proximity of valves, and the he lifted and displaced the extinguishers in order to update the data. Table 2 shows the data structure for the passive tags attached to fire extinguishers.

Table 3 Tag data structure for fire extinguishers

ID			DATA																				
			Specs	Status	Maintenance data															Environment Data			History
					Condition								Defective part							Location			
Type	Model	Serial	Manufacturing Date	Status	Obstructed	Pressure High	Pressure Low	Overall	Loose	Dusted	Rusted	Damaged	Missing Pin	Missing Rivet	Missing Paper	Missing Sign	Neck Bended	Plugged Hose	Seal Broken	Building	Floor	Room	Last Inspection Date

Due to the limited memory of passive tags, the above information is squeezed to binary codes and stored on the tags. The software translates BIM data related to components to codes using lookup tables and store codes in designated memory spaces on the tags.

The information about the defective parts that is written on the tag helps maintenance workers to quickly identify the problem based on the previous inspection and decreases the re-work. The user interface provides wizards for the inspector which contains standard instructions for inspection, alerts that are triggered by data read form the tags and customized data entry dialogue boxes based on the type of component.

The software also provides navigation aid for the inspector to locate the extinguishers and valves in the building using active tags. The software has pre-loaded floor plans as a visualization layer. By surveying the area to detect the tags, the sensed tags are shown on the floor plan based on their location information.

This case study has been done in a pilot scale in EV building of Concordia University where active and passive tags were attached to 9th floor fire valves and extinguishers. The technological feasibility of the system has been tested in a real working environment.

Conclusions and Future Work

The proposed methodology provides conceptual data structure and implementation approach of a futuristic vision of facilities with RFID tags attached to their components. Although the case studies show the technical feasibility of the proposed framework using available hardware, several challenges should be addressed to make the vision practical and financially feasible.

The following steps are necessary for realizing the proposed approach: (1) identifying most suitable building components for tagging based on cost-benefit analysis considering long-term value adding benefits, (2) re-engineering existing construction and maintenance processes for the selected components, (3) investigating product-specific and detailed tag structure for the selected components, (4) extracting important *process data* to be stored on the tags for each lifecycle stage of selected components, (5) technology selection and field testing for available RFID hardware, and (6) investigating new information to be added to BIM related to RFID.

Acknowledgement

The authors would like to acknowledge the contributions of all students who helped in realizing the case studies. Data collection and execution of case studies could not be done without the information provided by Concordia University Facilities Management Office and Environmental Health and Safety Office.

References

[1] Aimglobal (2008). "What is RFID", Retrieved October 2008, from: <http://www.aimglobal.org/technologies/rfid/what_is_rfid.asp>.
[2] East, W. E., and Brodt, W. (2007). "BIM for Construction Handover", JBIM, 28-35.
[3] Ergen, E., Akinci, B., and Sacks R. (2007). "Life-cycle data management of ETO components using radio frequency identification", Advanced Engineering Informatics, 21(4), 356-366.
[4] Harrison, M., and Parlikad, A. K. (2006). "Lifecycle ID and lifecycle data management", AUTO-ID Labs, AEROID-CAM-005.
[5] Isikdag, U., Aouad, G., Underwood J., and Wu, S. (2007). "Building information models: A review on storage and exchange mechanisms", Bringing ITC Knowledge to Work, 24th W78 Conference.
[6] Motamedi, A., and Hammad, A. (2009). "Lifecycle management of facilities components using radio frequency identification and building information model", Next Generation Construction IT, journal of IT in construction (Paper accepted).
[7] Song, J., Haas, C. T., Caldas, C., Ergen, E., and Akinci, B. (2006). "Automating the task of tracking the delivery and receipt of fabricated pipe spools in industrial projects", Automation in Construction, 15(2), 166-177.
[8] Stephens, J., Leonard, D., Groom, C., and Nisbet, N. (2005). "Life-cycle data for buildings: Opportunities for the IFC model based operation and maintenance (Ifc-mBomb)", IAI UK Ltd.

Development of Object Detection Technology Using Laser Sensor for Intelligent Excavation Work

Ji-Yune Soh[1], Junbok Lee[2] and Choong-Hee Han[3]

[1]Ph.D. Candidate, Dept. of Architectural Engineering, Kyung Hee University, Seochun-dong, Gihung-gu, Yongin-si, Korea ; PH +82-31-201-2966 ; FAX +82-31-204-6780 ; e-mail : kellysoh@khu.ac.kr
[2] Associate Professor, Dept. of Architectural Engineering, Kyung Hee University, Seochun-dong, Gihung-gu, Yongin-si, Korea ; e-mail : leejb@khu.ac.kr (corresponding author)
[3] Professor, Dept. of Architectural Engineering, Kyung Hee University, Seochun-dong, Gihung-gu, Yongin-si, Korea ; e-mail : chhan@khu.ac.kr

Abstract

The demand for the development of intelligent construction equipments is increasing gradually to deal with the current problems of construction technology such as lack of skilled workers, aging of engineers, safety issues, etc. Especially, earth work such as excavation is very machine-dependent, and there have been many researches on the development of intelligent excavator. This requires a great safety concern. Thus, the objective of this study is to develop the technology of enhancing the safety of intelligent excavation system by developing the essential technology of object detection for the intelligent excavator and applying it to a user-friendly system. Literature was reviewed, and the function of various sensor technologies was investigated and analyzed. Then, the best laser sensor was selected for an experiment to determine its effectiveness. Object detection algorithm was developed for a user interface program, and this can be used as the fundamental technology for the development of a safety management system for intelligent excavation work.

Keywords: Sensing, Laser Sensor, Backhoe Excavator, Automation, Safety Management

1. Introduction

Recent Korean construction industry is following the trend of state-of-the-art facility, complexity and gentrification in response to various demands of the consumers, and the attainment of the essential technology has been the key to keeping the competitive edge. Additionally, the generation of high added-value has been the objective of the construction industry to maintain the competitiveness in response to free market influence of construction industry. Nevertheless, construction industry has been traditionally regarded as one of 3D (dirty, difficult, dangerous) work, and the retention of young experienced workers has been a big problem due to the avoidance of young workers delving into the construction work. Old workers, who have been the main workhorse for the modernization in the past, have become aged, while the ratio of young workers desiring to be the main workhorse for future construction market has decreased considerably. Accordingly, the imbalance of supply and demand for construction workers has become a serious issue, and, moreover, industrial accidents have been reported the most in the construction industry. Figure 1 illustrates this point.

These problems of lack and aging of skilled workers and safety problems lead to the reduction of worker's productivity, worsening of the payback on investment due to increased wage, difficulty in maintaining the quality and safety as well as decreased competitiveness in construction technology. These are the current and impending issues for domestic construction industry. Accordingly, the development of automated construction equipments can be regarded as the best technological approach to solve these impending problems.

This study aims at developing an object detection technology for obstacles around the excavator using laser sensor in order to enhance the safety of the earth work environment. Literature was reviewed, and the function of various sensor technologies were investigated and analyzed to identify and select the best sensor.

Object detection algorithm was developed, and this was applied to safety management system for efficient operation and management of construction equipments.

Since the development of automated construction machine is in the initial stage, this study limited its scope to automated control of intelligent excavator and excluded working environment at inclined surface so as to consider flat (even) surface working environment only. Thus, the scope of this research for the development of an object detection technology for the safety management system of an automated excavator is so limited accordingly.

This study aimed at verifying the object detection performance of a laser sensor, selected as an essential technology. As the preliminary step prior to fully-installed experiment, only functional test and field test were carried out.

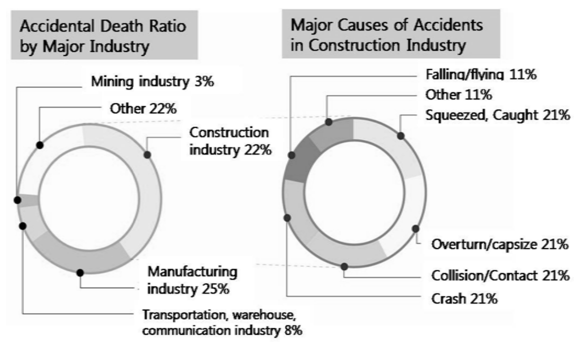


Figure 1 Status of Accidents in Construction Industry (The Korea Occupational Safety & Health Agency, 2006)

2. Core Technology for Object Detection

2.1 Selection of Core Technology

Various conditions applicable to an automated excavator were identified and analyzed for their application to each essential technology.

As a condition required for object detection technology applicable to safety management system, real-time detection must be possible, first of all, and the distance measurement between the object and excavator should be feasible so that the distance to the object can be measured and the collision to the object can be avoided. Additionally, it should be operable during movement or vibration owing to the environmental characteristics, which requires the object detection device to remain attached to the automated excavator.

As seen in table 1, which compared various technologies, sonar sensor and laser sensor were found to be most suitable for object detection.

Sonar sensor of the two previously selected sensors is sensitive to wind effect, and it is also subject to high risk of operation failure during bad weather. Thus, it is difficult to apply it to detection of objects in long distance at earth work field. In contrast, laser sensor is relatively more expansive but is faster in data processing. Moreover, it can detect objects in long distance and at earth work field. Thus, laser sensor was finally chosen for the detection of objects around the excavation work site after comparing all various conditions for the object detection.Mobile robots do not move in a particular direction only, but they use a rotating scan sensor to cover wider area for their mobilization just like the rotating radar at an air traffic control tower. These equipments are widely used, and the laser sensor of SICK is used the most in outdoor environment.

The laser sensor is relatively more expensive, but some of its products are manufactured specifically for outdoor use. Its outdoor performance to scan and sense an object was proven through a field test. Figure 2 shows the result of field test.

Objects at long distance were easily detected, and sensing of objects during rain was also no problem. Left photo (Figure 2) is the actual image, and the image at right is a monitor screen shot, showing the movement of an object by the movement of a line centered on the location of a laser sensor installation toward the moving object.

The laser sensor produced by SICK of Germany was selected finally based on the comparison analysis of economy, stability, application cases, etc. Table 2 shows major specifications of the selected sensor.

Table 1 Comparison of the applicability of essential technologies to object detection

	Real-time Detection	omni-directional detection (360°)	distance measurement (more than 10m)	operability during movement or vibration
Laser Scan (Triangular)	unfeasible	unfeasible	feasible	unfeasible
Laser Scan (TOF)	unfeasible	feasible	feasible	unfeasible
Stereo Vision (Pattern matching)	feasible	unfeasible	feasible	feasible
Stereo Vision (Shape from Shading) (Structured Light)	feasible	unfeasible	feasible / requires light	feasible
Intelligent Camera	feasible	feasible	unfeasible	feasible
CCD Camera	feasible	unfeasible	unfeasible	feasible
Sonar Sensor	feasible	unfeasible	feasible	feasible
Laser Sensor	feasible	unfeasible	feasible	feasible

Figure 2 Object Detection by SICK Laser Sensor

2.2 Sensor Operation Algorithm

Figure 4 depicts the sensor activation algorithm for the detection of obstacle objects by the excavating robot, and it is largely divided into a sensor and sensor control system. The sensor is LMS-221 model of SICK and is operated by 24V battery.

When the excavator is being prepared, the sensor is activated at the same time to synchronize the activities. When the excavator begins to work, the sensor is activated so that it can detect any objects within the working environment of the excavator. When the sensor detects an object, it sends a warning signal and stops the excavating work at the same time so as to check for the object. Additionally, when an object is detected, an output of the information on the location of the object can be produced.

Table 2 illustrates the composition of the object detection system using laser sensor. The system can be largely classified into sensor part and control system part. The sensor part consists of the sensor and other devices to make the sensor operable. The control system consists of a computer to have the operation devices connected and operable and a software user interface program to verify object detection. Connection cables and other miscellaneous devices are not illustrated in this table.

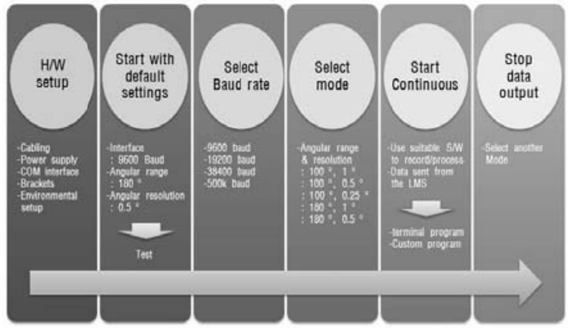
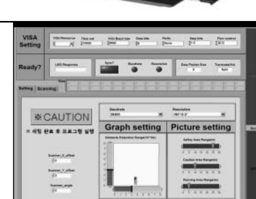


Figure 4 Sensor Activation Algorithms

Table 2 Composition of Object Detection System

<table>
<tr><td rowspan="5">Sensor</td><td>Laser Sensor</td><td></td></tr>
<tr><td>Power Supply</td><td></td></tr>
<tr><td>Brackets</td><td></td></tr>
<tr><td>24V Battery</td><td></td></tr>
<tr><td>Serial Hub</td><td></td></tr>
<tr><td rowspan="2">Control System</td><td>Hardware</td><td></td></tr>
<tr><td>Software</td><td></td></tr>
</table>

2.3 Sensor Detection Algorithm

Electricity is supplied to activate the sensor at the same time when the excavator is about to work. The sensor is subjected to an activation test with the values automatically assigned, i.e. the angle and speed of measurement and the allocated value of the angle. Then, the sensor begins to detect obstacles, and the software begins to execute its program. The initially allocated values can be changed.

3. Field Test

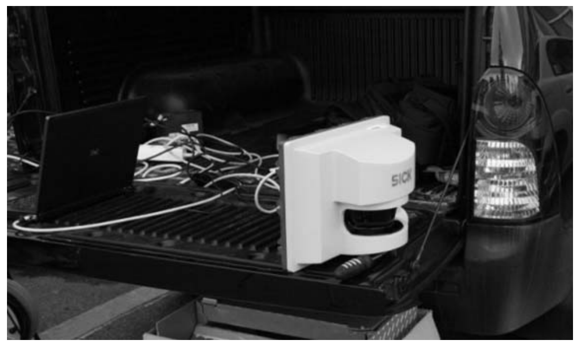


Figure 6 Image of sensor installation for the experiment

Figure 6 depicts the image of the sensor installed at the rear of an ordinary car for a field test. Its rear position was designed so that it can measure the outdoor environment characteristics and the performance during movement in the future experiment after its installation to the excavator.

3.1 Functional Test

Three functional tests were performed. The result of the tests indicated that obstacles at a long distance and moving objects could be detected, and this result could be extended to the earth work field site or even bad weather such as raining.

RS-232 was used as the data interface during field experiment, and the baud rate was set to 9600 for quick data acquisition. 180°/0.5° mode was used for wider measurement range and high resolution. This mode indicates the radius of 180° is sensed, and the detection continues at 0.5° interval.

The lower center of the graph is the location of sensor, and the sensor is set to cover 180° radius and up to 32m distance. When an object is not detected, the green line is marked as a semi-circle. However, when an object is detected, its shape changes by the location. As it can be seen in figures 7, 8, and 9, the place, where an automobile or a pedestrian passes by, the semi-circle shapes changes to a line, which moves toward the sensor depending on the distance.

Figure 7 depicts the result of first functional test. The first test involved the detection performance of the sensor, the distance to the object, and the kind of objects to be detected from the surrounding environment. The test result revealed that all surrounding objects including people, building, tree, stone, etc. could be detected. The place where an object such as an automobile is detected is marked by the circle to show the distance.

Figure 8 depicts the result of second functional test. The second test investigated the performance during raining and compared the output data of the distance and direction of the object in the sensing area with the actual measurement. The result indicated that the sensors operated well under bad weather conditions such as rain and wind. Additionally, the maximum distance error of 0.2m was observed from the measurement distance of 30m, while the direction of the object was accurately sensed.

Figure 9 illustrates the result of the third functional test. The third test involved the performance of the sensor in a very harsh environment of very bad weather and dusty earth work field site. The test result revealed that the sensors performed well at dusty earth work field site on a windy day to detect a person and construction materials scattered around the earth work field site.

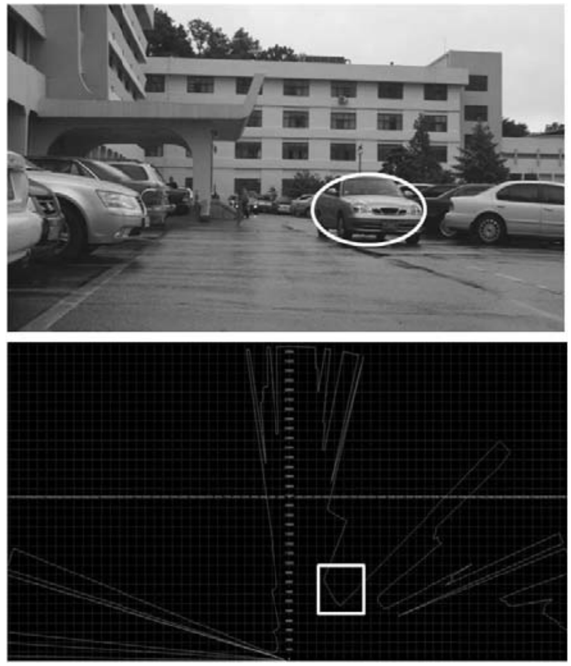


Figure 7 First Functional Test

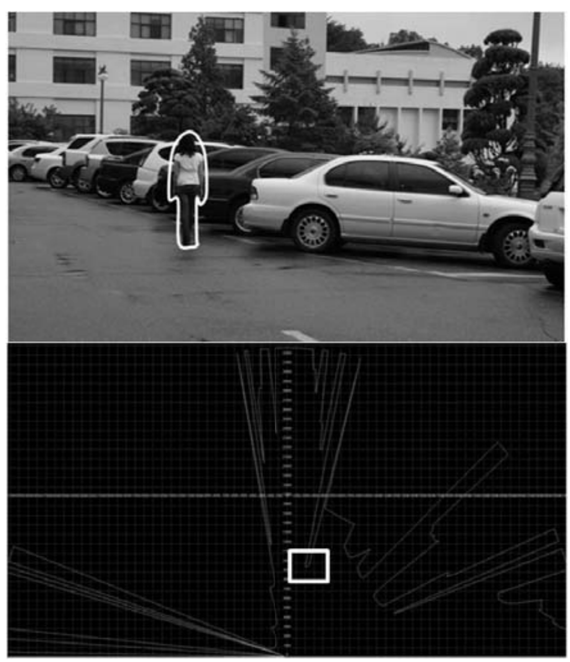


Figure 8 Second Functional Test

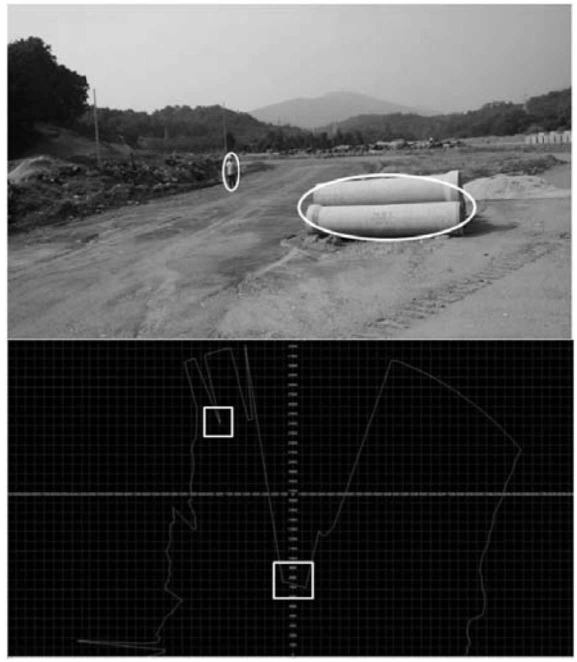


Figure 9 Third Functional Test

3.2 User Interface Performance Test

Three functional tests assessed the performance of the sensors, and two tests of user interface system were carried out. The results of the experiments indicate that the real-time detection of moving objects and obstacles were possible, and the information on the obstacles could be obtained from the user interface system. Additionally, the absolute coordinates, which were pre-set before the start of the sensor, made the detailed information on the direction and location of the objects possible. Additionally, a warning signal could be issued when an object is detected within a prescribed distance.

The bottom center of the semi-circle graph is the location of the sensor, and the sensors were designed to detect the objects at 180° range and up to 32m of distance. The area of detection of the objects is marked by one dot per 0.5° to indicate the location of the obstacle, and the dot is to move with the change in location of the object. Moreover, the dot changes its color depending on pre-specified distance, and a warning sound and message is issued when an object gets within a prescribed target distance. Figures 10 and 11 depict the place, where an automobile or pedestrian passes by, with gathered dots in the semi-circle to indicate the location, and change the color depending on the distance. Additionally, the accurate identification of the location can be possible through the screen conversion of the third grid background.

Figure 10 is the result of the first user interface operation. The test was carried out on still (not-moving) objects, and it was found that the location and distance of the objects were detected fairly well by the sensors. Additionally, the marks for obstacles changed the color at pre-specified distance interval, and a warning signal was issued right after the object appeared within a prescribed distance.

Figure 11 depicts the result of the second user interface operation. The second test was carried out on still objects and a moving obstacle, and the information on the location and distance of the obstacles detected by the sensors was fairly well represented by the user interface.

4. Conclusions

This study carried out a research on the development of object detection technology, which can be applied to earth work field site, as a preliminary investigation of the development of safety management system to be applied to automated excavation system. The following conclusions are derived from this research.

1) The validity of the core technology for object detection, i.e. laser sensor, selected from previous researches was affirmed through a field test.

2) A sensor operation algorithm was developed to apply the selected technology to the object detection system, and an object detection algorithm for the program to provide a user interface was also developed based on this sensor operation algorithm.

3) The developed algorithm was used to develop a user interface, using laser sensors, and the simultaneous operation of this user interface with the sensor was investigated through a field test.

This study carried out the development of essential technology and performance verification experiment only. It is expected and required to conduct additional experiments in the future so that the object detection system is to be installed on an excavator for the verification of the system applicability based on this real-time experiment.

The object detection technology and user interface, which were selected in this study, is expected to be used as the technology for an effective application of safety management system and for the attainment of real-time object detection data and database. Additional field experiments should be carried out to identify the location of the excavator deployment for an efficient application of these technologies. Moreover, if the blind spots can be eliminated with the all directions detection technology, then the developed object detection system can be commercialized.

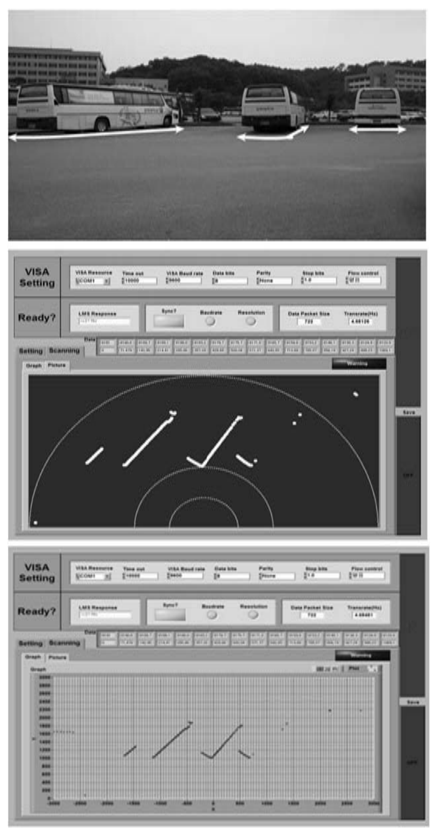


Figure 10 First User Interface Performance Test

5. Acknowledgement

This research was supported by a grant from Construction Technology Innovation Program (CTIP) funded by Ministry of Land, Transportation and Maritime Affairs (MLTM) of Korean government.

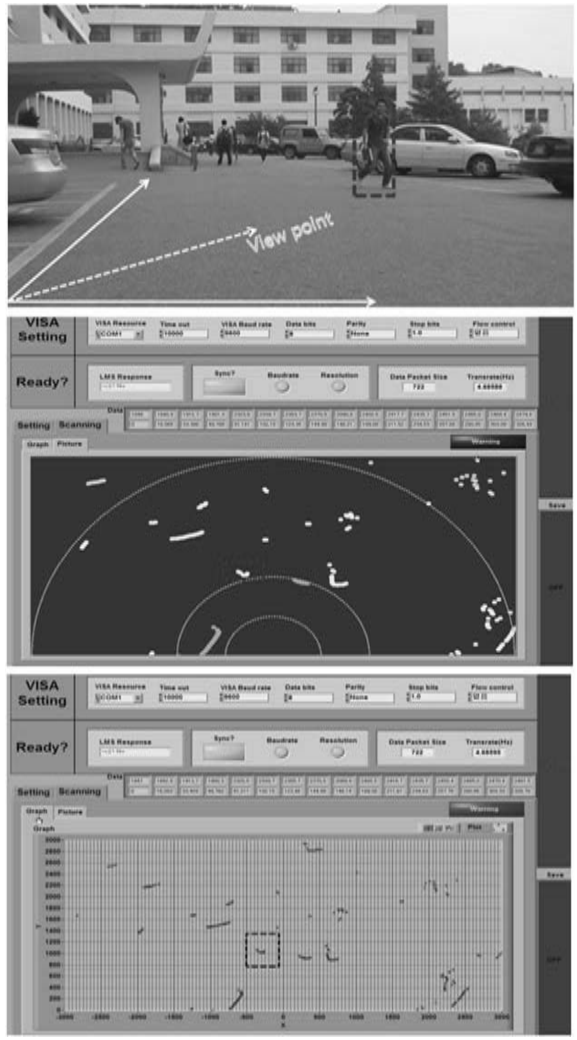


Figure 11 Second User Interface Performance Test

References

[1] Hyun J. Kang. (2005). "Real-time Object Detection and Tracing for Moving Objects." Journal of the Korea Society of Computer and Information.

[2] Seok M. Goh. (1986). "Investigation of Construction Automation System Technology." Korea Institute of Construction Technology.

[3] Jong P. Kim. (1992). "Investigation of Methods for Automation of Construction." Ministry of Education, Science and Technology.

[4] Young S. Kim. (2004). "Investigation of Essential technology for Automation of Construction and Strategy for Technology Renovation." Journal of Korean Society of Civil Engineers, 24/5D.

[5] Hwang Y. Yoo, et. al. (2007). "Development of Topographic Recognition Algorithm Using Sensor Fusion." RIST Research Paper.

[6] Electronic Engineer. (2006). "Automobile Vision System for Enhancing the Driver's Judgment."

[7] Anthony S Tentz. (2000). "A Robotic Excavator for Autonomous Truck Loading." The Robotics Institute, Carnegie Mellon University.

[8] Shigeru Sarata, Noriho Koyachi. (2007). "Detection of Dump Truck for Loading Operation by Wheel Loader." ISARC2007.

Research on Automated Project Performance Control: An Update

Ronie Navon

Associate Professor, Head, Construction Management & Economics – National Building Research Institute (NBRI), Faculty of Civil & Environmental Engineering, Technion City, 32000 Haifa, Israel. E-mail: ronie@technion.ac.il.

Introduction

Project performance control normally requires massive manual data collection – the more detailed the required control information, the more labor intensive the data collection must be. Hence, the cost of collecting the data and generating the information is high and the quality, integrity and real-time availability are low. Our proposition, for more than a decade, has been to either automate the data collection using advanced technologies, or use existing tools (e.g. daily site report – DSR) to minimize the data collection. This was done in a variety of areas: labor tracking; earthmoving operations; daily-site-report based control; change management tool; materials management and control; safety control; and company level monitoring – all are briefly reported in this extended abstract.

Labor tracking

Labor is one of the most significant resources in construction projects, hence controlling this resource can contribute meaningfully to the efficiency of managing the projects, both in terms of cost and progress. Our approach to labor tracking was to automate the data collection using automated location measurement technologies and to convert these locations into the desired information. The algorithm, developed for this conversion, automatically determines what activity is done at any given time (Goldschmidt and Navon 1996; Navon and Goldschmidt 1999; Navon and Goldschmidt 2003a; Navon and Goldschmidt 2003b). The algorithm was tested on site in a series of experiments to determine the feasibility of implementing it.

We are currently conducting simulated experiments of RFID technology to replace the location measurement technology. Instead of measuring the locations of the workers and associating them to building elements this algorithm determines which building element is worked on at each given time, based on RFID measurement.

Earthmoving Operations

The same principles, of automated location measurement and its conversion into control information, were used here. A model that uses location data measured with a GPS was developed and tested on site (Navon et al. 2004; Navon and Shpatnitsky 2005; Navon and Shpatnitsky 2006; Navon et al. 2002). This model required manual intervention at the end of each measurement day, which led to the next stage of the research: the dynamic work envelope (DWE) approach. In the DWE approach, instead of associating locations to activities by predetermining work sections, the algorithm determines the work envelopes dynamically during its operation, according to the measured locations. This approach is fully automated.

Daily-Site-Report Based control tool

The daily site report (DSR) logs most of the major events occurring in a construction site on a daily basis. The document is approved and signed both by the contractor and the owner. Hence, the DSR is a reliable and accurate source of data, capable of being available almost in real-time. In spite of this, it is very rarely used to control the project – it is mostly used for litigation purposes. The data in the DSR is still normally handwritten, which makes it hard to use even for litigation purposes because it is often illegible, unstructured and is not searchable.

We have developed a model, which uses an electronic format of the DSR, thus enabling the extraction of relevant data and their processing in order to generate real-time control information. The model provides an as-made schedule compared to the planned one (Navon and Haskaya 2007). In addition, the model calculates and compares the planned usage of resources to the ones actually consumed on site.

Change Management and Control Tool

Changes made in construction projects during their design and execution are a major cause for delays, cost overruns and deviations from performance requirements. The impact of changes often becomes clear only after their implementation in the project. At that stage it is difficult to make adjustments or consider alternatives. A timely recognition, by the project team, of the implications of proposed changes can lead to a reconsideration of the changes, so that the completed project will still meet the client's objectives.

We are developing a model that facilitates automatic identification of the possible consequences of changes when they are first proposed, prior to their implementation in the design and planning of the project (Isaac and Navon 2008; Isaac and Navon 2009). A graph-based approach is used in order to identify the impact of changes on the primary client objectives of cost, schedule and performance. Further research is being undertaken to develop the model's ability to quantify the extent of the change impact through the use of stochastic tools.

Materials Management and Control

Materials contribute a very significant percentage of the cost of construction projects. Workers are often idle, waiting for unavailable materials. Hence, efficient materials management and control can contribute not only to reduced materials waste, but also to increased productivity.

We have developed a model, which deals with materials purchasing, their delivery to the site and their dispatching for use in the building. The model generates reports and alerts – the reports include a list of all the materials needed in the project, a list of materials to be ordered, a cumulative list of materials flow, list of actually used materials vs. the planned ones and dead inventory, etc. The alerts include purchase orders (PO) not confirmed by the supplier, a list of materials that should have been ordered, but were not, a list of materials that were expected but were not supplied, materials arriving to the site which are incompatible with the PO, and the deviation between planned and actual quantities (Navon and Berkovich 2005; Navon and Berkovich 2006).

Safety Control

Even though fall from height is the number one risk factor in lethal accidents in construction, many of these accidents could have been avoided if the right preventive measures had been taken in time. The most common preventive measures are guardrails. The aim of our research was to develop an automated model that monitors guardrails in buildings under construction (Navon and Kolton 2006a; Navon and Kolton 2006b). The model identifies the activities associated with the risk of falling from heights and the areas where these activities are scheduled to be performed. Accordingly, the model plans the protective measures, namely the guardrails. The model follows up the existing guardrails and constantly compares their locations and lengths to the planned ones. Based on this comparison, the model issues warnings whenever guardrails are missing, or temporarily removed.

Company Level Monitoring

Financial ratios (FR) are useful indicators of a company's performance. Most ratios can be calculated from the information provided by the financial statements. FR can be used to analyze trends and to compare the company's financials to those of other companies. There are more than 25 FR, but their importance changes according to the type of industry, the characteristics of the company and the country that the company operates in. Limited research was done to identify the most relevant FR for the construction industry.

We interviewed financial managers of construction companies and identified five FR, which are most pertinent to construction companies: Net Income/ Equity; Equity/Total Assets; Operating Profit/Total Sales; Operating Profit/Interest Expense; and Financial Depth/EBITDA Net. As a result of the interviews, we also determined which values of the FR indicate a financially sound company ("desired values"). Then we analyzed the financial statements of 18 public construction companies and compared the actual values of their FR to the “desired values”. These “desired values” are probably more pertinent to the Israeli construction industry.

Conclusions

Our research, in a variety of areas, shows that data collection for real-time control can be reduced dramatically. As a result, the control information is less expensive to generate, more accurate, of a higher level of integrity and less error prone.

Acknowledgments

This research was partly supported by THE ISRAEL SCIENCE FOUNDATION (grant No. 444/05). This support is gratefully acknowledged. I would like to thank my colleagues for their invaluable contributions to the research described in this extended abstract. They are O. Berkovich, Y. Doytsher, E. Goldschmidt, I. Haskaya, S. Isaac, S. Khoury, O. Kolton, I. Paz-Chen, Y. Shpatnitsky, and A. Warszawski.

References

[1] Goldschmidt, E., and Navon, R. "Automated Real-Time Manpower Productivity Control." The Eighth International Symposium on Organization & Management of Construction, Glasgow, UK, 199-205.
[2] Isaac, S., and Navon, R. (2008). "Feasibility Study of an Automated Tool for Identifying the Implications of Changes in Construction Projects." Journal of Construction Engineering and Management, ASCE, 134(2): 139-145.
[3] Isaac, S., and Navon, R. (2009). "Modeling Construction Projects as a Basis for Change Control." Automation in Construction: (scheduled).
[4] Navon, R., and Berkovich, O. (2005). "Development and On Site Evaluation of an Automated Materials Management and Control Model." Journal of Construction Engineering and Management, ASCE, 131(2): 1328-1336.
[5] Navon, R., and Berkovich, O. (2006). "An Automated Model for Materials Management and Control." Construction Management and Economics, 24(6): 635-646.
[6] Navon, R., and Goldschmidt, E. "Automated Data Acquisition for On-Site Control." 16th IAARC/IFAC/IEEE International Symposium on Automation and Robotics in Construction (ISARC'99), Madrid (Spain), 681-686.
[7] Navon, R., and Goldschmidt, E. (2003a). "Can Labor Inputs be Measured and Controlled Automatically?" Journal of Construction Engineering and Management, ASCE, 129(4): 437-445.
[8] Navon, R., and Goldschmidt, E. (2003b). "Monitoring Labor Inputs: Automated-Data-Collection Model and Enabling Technologies." Automation in Construction, 12(2): 185-199.
[9] Navon, R., Goldschmidt, E., and Shpatnitsky, Y. (2004). "A Concept Proving Prototype of Automated Earthmoving Control." Automation in Construction, 13(2): 225-239.
[10] Navon, R., and Haskaya, I. (2007). "Is detailed progress monitoring possible without designated manual data collection?" Construction Management and Economics, 24(12): 1225-1229.
[11] Navon, R., and Kolton, O. (2006a). "Algorithms for Automated Monitoring and Control of Fall Hazards." Journal of Computing in Civil Engineering, ASCE, Submitted.
[12] Navon, R., and Kolton, O. (2006b). "A Model for Automated Monitoring of Fall Hazards in Building Construction." Journal of Construction Engineering and Management, ASCE, Accepted.
[13] Navon, R., and Shpatnitsky, Y. (2005). "Field Experiments in Automated Monitoring of Road Construction." Journal of Construction Engineering and Management, ASCE, 131(4): 487-493.
[14] Navon, R., and Shpatnitsky, Y. (2006). "A Model for Automated Monitoring of Road Construction." Construction Management and Economics, 23(9): 941-951.
[15] Navon, R., Shpatnitsky, Y., and Goldschmidt, E. "Model for Automated Road-Construction Control." Nineteenth International Symposium on Automation and Robotics in Construction, 425-430.

3D Measuring and Marking System for Building Equipment: Developing and Evaluating Prototypes

Shintaro Sakamoto, Hiroki Kishimoto, Kouetu Tanaka and Yukiteru Maeda

Research and Development Center, Shinryo Corporation, 41 Wadai, Tsukuba, Ibaraki, 300-4247, Japan,
PH 81 298 64 6105; Fax 81 298 64 6127; e-mail:sakamoto.sh@shinryo.com

Abstract

The measuring work has been done manually with the scales in the construction of building equipments. We have been studying a 3D measuring and marking system using a total station to increase accuracy and decrease workload. One feature is the pointing device for measuring indirectly the hidden areas which cannot be measured with a stand-alone total station. The device has a sensor which consists of a pinhole and a position sensitive detector for measuring the incident angles of the received laser and it's possible to calculate the pointing position by using the data. The other is the positioning device for marking accurately relative to the points which the total station indicates with the laser automatically. The displacement from the target position is calculated from the device position and the operator marks with the marking position adjusting function of the device. This paper describes the system configuration, measuring and marking theory of the system with those devices, and the results of evaluation experiments with prototypes.

Keywords: Building Equipment, 3D Position Measuring and Marking System, Total Station, Position Sensitive Detector

1. Introduction

In the construction of air-conditioning systems, the position of system elements is marked out on structures such as floors and ceilings by measuring the distance from the datum line of the building with reference to the drawings. In renewal projects, construction planning is based on the measurement data of the existing equipment. Conventionally, these measurements represent two-dimensional distance data obtained using a scale or a compact laser range finder. It is considered that a shift to 3D coordinate measurement and seamless connection of a 3D CAD system and on-site measuring system will significantly improve the accuracy of measuring and the efficiency of work, concretely, decreasing the rework substantially.

Indoor 3D measuring methods based on the principle of GPS have been proposed in the past but these have lacked sufficient accuracy for position measurement in construction applications. Additionally, they are inadequate for marking applications because they cannot locate target positions directly to the site. In more recent years, instruments that measure 3D coordinates by measuring distance and angle, such as laser scanners and total stations, have been used in construction fields. But because they need to be moved to measure the positions in hidden areas, they tend to increase the workload and lower accuracy through the frequent displacement, because most points of measurement in equipment work are in a crowded location, a machine room, for example. And a laser scanner is impractical for acquiring measuring data on-site because of the specialized technique required and because several million sets of point data have to be processed. In marking work, compensation for the irradiating position is necessary because the positions depend on the accuracy of structures such as the floor, walls and ceiling. In addition, laser spot stretching because of slanting decreases marking accuracy.

In this research, we have developed a 3D measuring system for equipment work. The system consists of a total station, a pointing device for indirect measurement of the position in hidden areas and a positioning device for improving marking accuracy. This paper will show the configuration, system usage, and the mechanism and measuring principles of each device. Also, it will show the results of evaluation experiments using a prototype of the system.

2. System Configuration and Modes of Operation

This system is based on the total station concept and comprises a pointing device, positioning device and a host PC. The total station is a Leica Geosystems TCRP1205+. The accuracy of angle measurements is 5 seconds and distance accuracy is 1 mm + 1.5 ppm in the prism mode and 2 mm + 2 ppm in the non-prism mode.

This system has three operational modes: an initializing mode, position measurement mode, and positioning mode. The initializing mode is used to measure the self-position of the total station with reference to the datum line of the building or two points whose coordinates are known. The reference points are measured by setting a prism on them. The position measurement mode is used for measuring the 3D coordinates of not only the point collimated directly but also the hidden point with the pointing device. In the case of using the device, the operator establishes contact between the tip of the device and the objective position, and then adjusts the device posture as the laser is irradiated to the appropriate position. The total station measures the device position and the device measures the laser incident angles and the rotational angle around the laser. The system calculates the 3D coordinates of the pointed position with those measurement values. In the positioning mode, the total station irradiates the laser to the target position input to the system preliminarily or on site. After setting the positioning device on the position, the system measures the distance and the angles to the prism on the device. The displacement from the target position is calculated and shown to the operator. The operator marks referring the value.

3. Pointing Device

3.1 Angle Parameter of Device

As shown in Figure 1, the position of the laser receiving point P0 is measured directly on Σ_1 whose origin is the position of the Total Station. The posture of the device is determined by the rotational angle around the laser axis and the laser incident angles. The local coordinate system of the device Σ_2 is defined as the origin is P0, the z axis is conformed to the laser direction and the x axis is on the vertical plane. The rotational angle around the z axis is defined as γ, the angle of the laser incident direction is α and the angle between the laser and the normal vector of the receiving surface is β on Σ_2 as shown in Figure 2. The coordinates of the pointing position are calculated by these three angles.

3.2 Laser Incident Angle Sensor

The device has an angle sensor composed of the pinhole and position-sensitive detector or PSD, for measuring the angle of α and β. PSD is the sensor which detects the position of the light spot via the surface resistance of the photodiode, a 2D PSD "S1880" produced by Hamamatsu Photonics, shown in Figure 3. As shown in Figure 4, when the laser is irradiated to the pinhole, the laser which passes through the pinhole is irradiated to the PSD as laser spot S. Because its position depends on the angle of α and β, it is possible to measure them from the position of S. Here, α is defined as the angle from the x axis of PSD and calculated using the following equation:

$\cos\alpha = x_p / r$, where S = S(xp,yp) and $r = \sqrt{{x_p}^2 + {y_p}^2}$ (Eq. 1)

β is expressed as the function of the length r shown in equation 1 and is constantly plus. The incident angle sensor is a modular structure and is demountable from the device, enabling it to be set to calibration equipment as shown in Figure 5. Using this equipment, the calibration curve of β is obtained from data r and the inclination angle of the sensor module by a least square method, as shown in Figure 6. Also, the offset values depending on the mechanical accuracy of PSD and its embedding are measuring with this equipment.

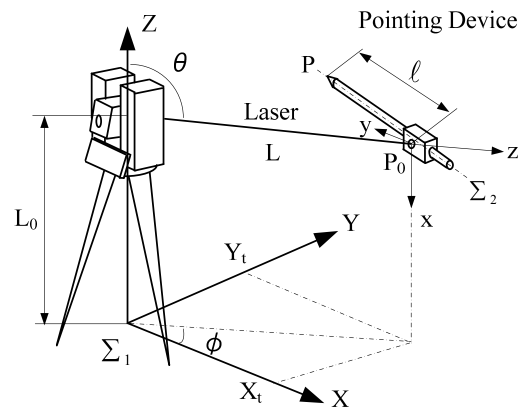


Figure 1 Definition of Coordinates

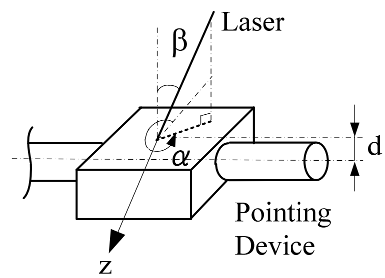


Figure 2 Definition of α, β

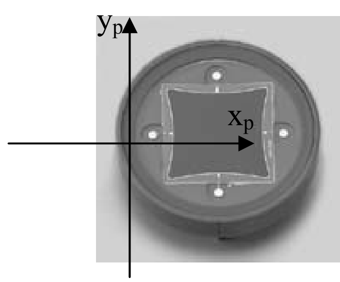


Figure 3 Position Sensitive Detector (PSD)

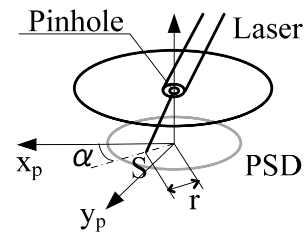


Figure 4 Measurement of α, β

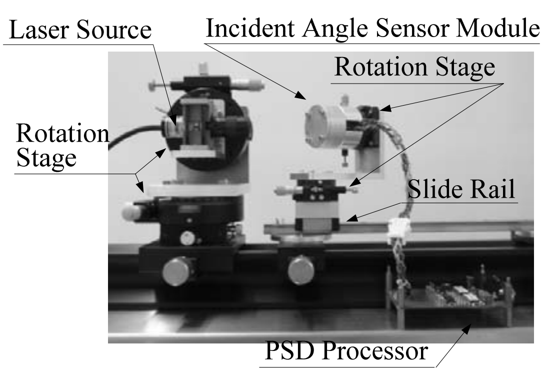


Figure 5 Calibration Equipment

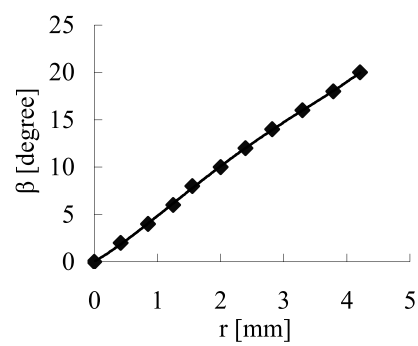


Figure6 Calibration Curve of β

3.3 Measurement of γ

The rotational angle γ around the z axis on the device coordinates system is measured using the inclinometer embedded parallel to the laser receiving surface. It enables angle measurement of 360 degrees. The resolution is 0.00549 degree and the linearity $\pm$0.2 degree. However, γ cannot be measured directly by the sensor because the sensor's measurement value is affected by α, β and the vertical angle of P0 θ, shown in Figure 1. Therefore, γ is measured indirectly using the fact that the angle æ between the vector $\vec{M}$, meaning the basic axis of the inclinometer, which is transformed by α, β, γ and θ, and the vector $\vec{G}$, meaning the axis of the gravity orthographically projected to the laser receiving surface, are equal to the measurement value of the inclinometer. The angle æ is found by the following equation.

$$\cos \psi = \vec{M} \cdot \vec{G} / \|\vec{M}\| \cdot \|\vec{G}\| \quad \text{(Eq. 2)}$$

The vector $\vec{M}$ is obtained by translating the vector of the basic axis of the inclinometer on Σ_2, $\vec{m} = [1,0,0]^T$, as the following. Here, the trigonometric function of cosine and sine are shown in the abbreviation as C and S.

(1) Rotation by γ

$$\vec{m}' = \begin{bmatrix} C_\gamma & -S_\gamma & 0 \\ S_\gamma & C_\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \vec{m} = \begin{bmatrix} C_\gamma \\ S_\gamma \\ 0 \end{bmatrix} \quad \text{(Eq. 3)}$$

(2) Transform by and

As shown in Figure 7, the transform by and is equal to the transform of rotating in around the vector $\vec{\omega} = [S(\alpha-\gamma), C(\alpha-\gamma), 0]^T$.

$$\vec{m}'' = (\vec{m}' \cdot \vec{\omega}) \cdot \vec{\omega} + C_\beta \{\vec{m}' - (\vec{m}' \cdot \vec{\omega}) \cdot \vec{\omega} + S_\beta (\vec{\omega} \times \vec{m}')\} = \begin{bmatrix} S_\alpha S_{\alpha-\gamma} + C_\beta (C_\gamma - S_\alpha S_{\alpha-\gamma}) \\ S_\alpha C_{\alpha-\gamma} + C_\beta (S_\gamma - S_\alpha C_{\alpha-\gamma}) \\ -C_\alpha S_\beta \end{bmatrix} \quad \text{(Eq. 4)}$$

(3) Rotation by θ

$$\vec{M} = \begin{bmatrix} C_\theta & 0 & S_\theta \\ 0 & 1 & 0 \\ -S_\theta & 0 & C_\theta \end{bmatrix} \cdot \vec{m}'' = \begin{bmatrix} C_\theta \{S_\alpha S_{\alpha-\gamma} + C_\beta (C_\gamma - S_\alpha S_{\alpha-\gamma})\} - S_\theta C_\alpha S_\beta \\ S_\alpha C_{\alpha-\gamma} + C_\beta (S_\gamma - S_\alpha C_{\alpha-\gamma}) \\ -S_\theta \{S_\alpha S_{\alpha-\gamma} + C_\beta (C_\gamma - S_\alpha S_{\alpha-\gamma})\} - C_\theta C_\alpha S_\beta \end{bmatrix} \quad \text{(Eq. 5)}$$

The vector $\vec{G}$ is obtained by translating the gravity vector $\vec{g} = [0,0,-1]^T$ with the orthographical projection matrix Q derived from the direction cosine of the receiving surface. The direction cosine is derived by the same translation of the vector $\vec{n} = [0,0,1]^T$ as the vector $\vec{M}$ with equations 1, 2 and 3. The vector $\vec{G}$ is expressed as equation 6.

$$\vec{G} = \mathbf{Q} \cdot \vec{g} = \begin{bmatrix} (C_\theta S_\beta C_{\alpha-\gamma} + S_\theta C_\beta) \cdot (C_\theta C_\beta - S_\theta S_\beta C_{\alpha-\gamma}) \\ -S_\beta S_{\alpha-\gamma} (C_\theta C_\beta - S_\theta S_\beta C_{\alpha-\gamma}) \\ (C_\theta C_\beta - S_\theta S_\beta C_{\alpha-\gamma})^2 - 1 \end{bmatrix} \quad \text{(Eq. 6)}$$

3.4 Derivation of Pointing Coordinate

As shown in Figures 1 and 2, the Total Station height is described as L0, the distance to the receiving surface is L, the horizontal angle is ϕ and the vertical angle is θ on Σ_1. When the axial length between P0 and P is described as ℓ and the offset of the thickness direction is d, the pointing position P[0, ℓ ,d] is translated to P" with equations 3 and 4.

$$P'' = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} = \begin{bmatrix} C_\alpha S_{\alpha-\gamma} \ell - C_\beta (S_\gamma + C_\alpha S_{\alpha-\gamma}) \ell + S_\beta C_{\alpha-\gamma} d \\ C_\alpha C_{\alpha-\gamma} \ell - C_\beta (C_\gamma - C_\alpha C_{\alpha-\gamma}) \ell + S_\beta S_{\alpha-\gamma} d \\ C_\beta d + S_\alpha S_\beta \ell \end{bmatrix} \quad \text{(Eq. 7)}$$

The pointing position Pt is derived by translating P" to Σ_1.

$$P_t = \begin{bmatrix} C_\phi & -S_\phi & 0 \\ S_\phi & C_\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} C_\theta & 0 & S_\theta \\ 0 & 1 & 0 \\ -S_\theta & 0 & C_\theta \end{bmatrix} \cdot \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} + \begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} C_\phi (C_\theta P_1 + S_\theta P_3) - S_\phi P_2 + S_\theta C_\phi L \\ S_\phi (C_\theta P_1 + S_\theta P_3) + C_\phi P_2 + S_\theta S_\phi L \\ -S_\theta P_1 + C_\theta P_3 + C_\theta L + L_0 \end{bmatrix} \quad \text{(Eq. 5)}$$

Here, the coordinates of P0 are defined as [Xt,Yt,Zt]T. Finally, Pt is translated to the global coordinate

by using the position of the Total Station and the rotation angle around the axis Z0 measured in the initialize mode.

3.5 Device Configuration

The prototype of the pointing device is shown in Figure 8. The length of ℓ is 1000 millimeters and the ribs are assembled on the bar to suppress deflection. The microprocessor calculates the three angles mentioned above with the output of the sensors. It also sends control commands (device search and measurement) to the host PC via Bluetooth modem, and measurement data are sent between the two sides.

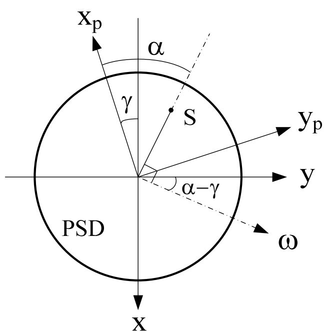


Figure 7 Transform by α and

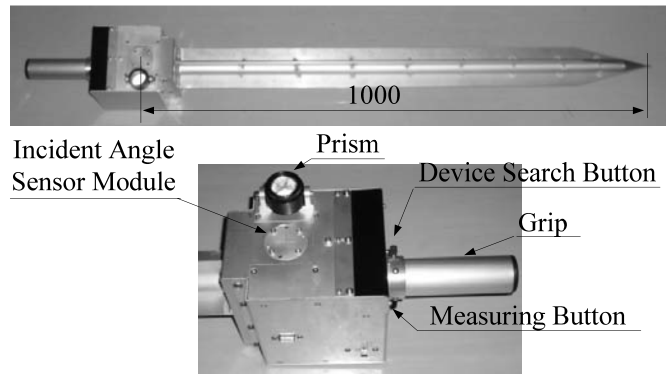


Figure 8 Pointing Device

The prism is mounted on the receiving surface to be found by the total station.

4. Positioning Device

4.1 Theory of Positioning

This system has a positioning device for accurate positioning and marking. We produced a prototype for use on the floor and on machine foundations. As shown in Figure 9, the laser is irradiated to a different point P' from the target P because of the level difference in the case of the positioning on the floor. In use, the operator sets the device after irradiation and measures the horizontal distance L0, from the total station to the prism. The operator measures the horizontal distance L1 from the prism to the target point and then measures L1 with the device and marks.

4.2 Device Configuration

The prototype of the positioning device is shown in Figure 10. It consists of the prism, axial slider, scale, slit, screws and level. The level of the device is adjusted via the screws and the axial direction via referring to the laser through the slit and the vertical line on the slider. The slider is moved along the scale for positioning and the operator makes marking with a pen through the hole on the slider.

5. Results of Experiments

5.1 Conditions of Experiments

We conducted experiments to evaluate the measuring accuracy of the prototype system. The experimental field is about 5 meters wide and 12 meters long. It is an indoor field and is not subjected to direct sunlight. The reference values for evaluation are measured with the two transits set on the reference points. The transit is NET 2000 products produced by SOKIA and the angle resolution is 2 seconds. The target accuracy setting of this system is 5 millimeters at 10 meters from the total station.

5.2 Results of Experiments

5.2.1 Position Measurement Mode

The measurement points P1 and P2 are set on the floor. Here, P1 and P2 are defined on each coordinate system, and as shown in Figure 11. S1 and S2 are the reference points of each coordinate system and the position of the two transits. The position of the total station is measured by the transits as T1(2399.5, 8384.3, 1424.8) on Σ_{01} and T2(2340.7, 5082.1, 1425.1) on Σ_{02}. The evaluation is done in six cases of the device posture as shown in Figure 12. The posture which is set on the measurement points at reasonable angle is defined as the default, and it's changed to the six cases while the tip is kept to set on the points.

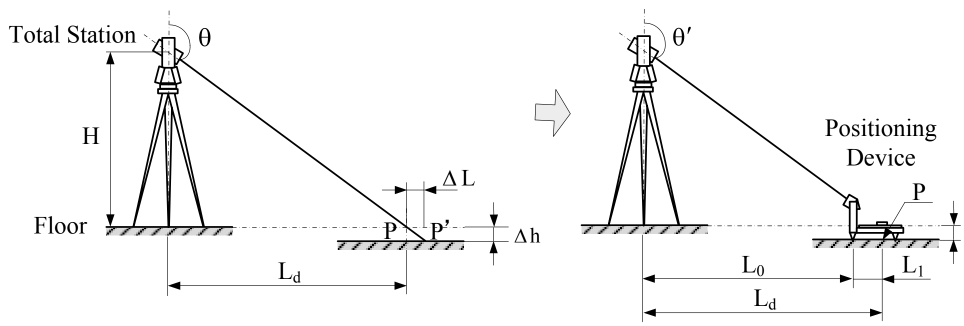


Figure 9 Theory of Positioning

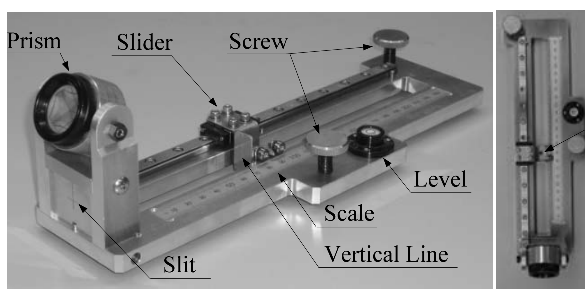


Figure 10 Positioning Device

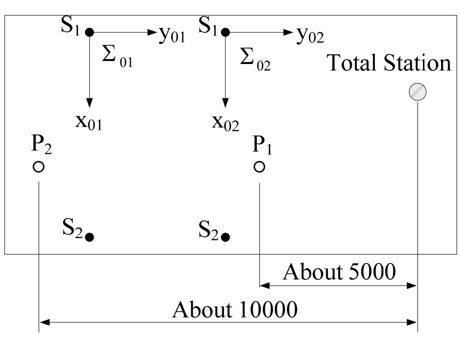


Figure 11 Experiment Field and Setting

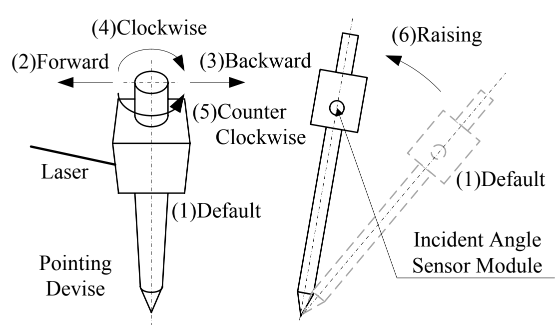


Figure 12 Posture of Pointing Device of Measurement Mode

The results are shown in Table 1. The measurement errors are evaluated as the differences from the results measured by the transits and expressed as the root sum square of each coordinate error. They are less than 5 millimeters on both points except for the one case of (6) at P1. It is supposed that the error is from the measurement accuracy of the sensors, especially the incident angle sensor module. It is considered to be necessary to analyze laser behavior in the incident angle sensor module and to improve the calibration equipment for accuracy and repeatability.

6. Conclusions

We have proposed a 3D measuring system for application in the construction field, especially for equipment work. In this paper, we presented a prototype of the system and the results of the experiments to evaluate the accuracy of measurements. The results of experiments demonstrated that the accuracy of position measurement with the pointing device is approximately 5 millimeters and the accuracy of positioning with the positioning device is approximately 2 millimeters. In the future, we will analyze the laser behavior and reconsider the configuration of the incident angle sensor and the method of the calibration to improve the accuracy of measuring with the pointing device. We will also redesign it for reducing the size and weight, and consider the other pointing methods such as the laser. Concerning the positioning device, we will make improvements so that it can be used on walls and ceilings.

Table 1 Results of Point Measurement Mode

			X[mm]	Y[mm]	Z[mm]	Error[mm]
P_1	Result by Transits		3192.8	3539.9	5.5	
	Result by the System at Each of the Six Postures	(1)	3192.3	3542.2	4.7	2.5
		(2)	3192.0	3537.6	4.9	2.5
		(3)	3192.3	3543.1	5.0	3.3
		(4)	3192.3	3537.4	4.9	2.6
		(5)	3192.4	3538.9	4.9	1.2
		(6)	3191.9	3545.8	3.7	6.2
P_2	Result by Transits		3262.8	-4762.0	11.9	
	Result by the System at Each of the Six Postures	(1)	3262.7	-4761.3	10.7	1.4
		(2)	3264.0	-4760.0	10.3	2.8
		(3)	3263.2	-4760.2	11.0	2.1
		(4)	3262.4	-4766.6	11.0	4.7
		(5)	3261.6	-4763.1	9.0	3.3
		(6)	3262.2	-4761.3	10.6	1.6

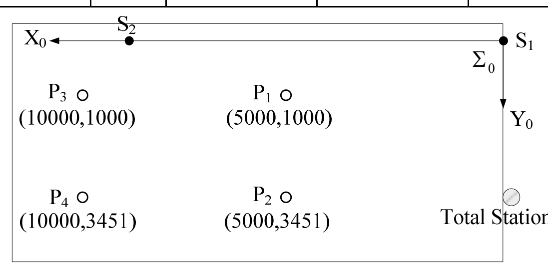


Figure 13 Experiment Field and Setting of Positioning Mode

Table 2 Results of Positioning Mode

Point	No.	X [mm]	Y [mm]	Error [mm]		
				E_X	E_Y	RSS
P_1	1	4998.7	1000.1	-1.3	0.1	1.3
	2	4999.3	999.8	-0.7	-0.2	0.7
	3	4999.3	1000.3	-0.7	0.3	0.8
P_2	1	5000.0	3450.6	0.0	-0.4	0.4
	2	5000.1	3450.9	0.1	-0.1	0.1
	3	4999.9	3450.7	-0.1	-0.3	0.3
P_3	1	9998.9	998.1	-1.1	-1.9	2.2
	2	9998.9	998.1	-1.1	-1.9	2.2
	3	9998.8	998.2	-1.2	-1.8	2.2
P_4	1	10000.0	3448.9	0.0	-2.1	2.1
	2	10000.1	3448.8	0.1	-2.2	2.2
	3	10000.0	3449.1	0.0	-1.9	1.9

References

[1] Ward, A., and et al. (1997). "A New Location Technique for the Active Office", IEEE Personal Communications, Vol.4, No.5, pp.42-47.

[2] Kien, D. T., and et al. (2002). "Un ultrasonic Indoor Positioning System with Self-Configuration", Proceeding of IEICE General Conference, pp.243

[3] Pedrovski, I., et al., (2004) "Pseudolite Application for ITS", Technical Report of IEICE, pp.13-18.

[4] Kawamura, K., (2006) "Engineering Process Innovation by 3D Laser Scan", Journal of Japan Society for Design Engineering, pp.83-87.

[5] Ohtsu, S., et al., (2007) "Method of processing data with a cloud of points on the 3-D shape surveying", Journal of applied computing in civil engineering, Vol.16, pp. 27-36

Time and Cost Evaluation of Construction of Steel Framed Composite Floor with Precast Concrete Floor Structure

U. D. Dabhade[1], N.A.Hedaoo[2], Dr. L. M. Gupta[3] and Dr. G. N. Ronghe[4]

[1]M.Tech. (C.M) Student, Department of Civil Engineering, College of Engineering, Pune-411005, India; PH 091-020-25507090; FAX 091-020-25507299;e-mail: u_dabhade@yahoo.co.in
[2]Sr.Lecturer, Department of Civil Engineering, Govt. College of Engineering, Pune-411005, India; PH 091-020-25507090 FAX 091-020-25507299; e-mail: nahedaoo@yahoo.com
[3, 4] Professor, Department of Applied Mechanics, Visvesvaraya National Institute of Technology, Nagpur-11, india; PH 091-0712-2222828; FAX 091-020-2223230; e-mail: lmgupta_vrce@yahoo.co.in and e-mail: gnronghe@indiatimes.com

Abstract

The present scenario in India, particularly in metro cities has restricted the horizontal growth, which led to the vertical growth for building construction. Today, fast track construction is a rapidly growing economy, brings rising costs and therefore time saving in construction can compensate significant proportions of the overall construction cost.

This paper presents a study on, time and cost wise feasibility of steel framed composite floor building. A case study considered for this work is 10 storied multilevel cars parking building. A major feature of this building is post-tensioned composite steel beams having span of 16 m. Considering same plan, floor area, floor to floor height and loading conditions, this existing building is designed and constructed by other two ways viz. precast concrete frame with precast concrete floor and steel frame with precast concrete floor. While designing the above structure with precast concrete frame with precast concrete floor, one additional column is introduced in between 16m span lengths to the overall plan to suit the design criterion.

The Microsoft Project-2003 used for time scheduling and the optimum time required at different stages of all these three buildings are calculated. The total cost of each structure is calculated as material and construction costs of each structural element only. The results shows that steel frame with composite deck floor saves 55.3% construction time than precast frame with precast concrete floor and 14.3% compared to steel frame with precast concrete floor. However, this required extra 23.10% of direct cost and 12.99% of net cost for precast frame with precast concrete floor while 0.52% and -2.34% for steel frame with precast concrete floor.

Keywords: Composite floor construction; Post-tensioned composite beam; Precast Concrete floor; Time scheduling; Cost evaluation;

1. Introduction

One of the biggest revolutions came with introduction of hot-rolled steel section and cold-formed steel decking as a construction material for high-rise buildings. Steel framed structures with the composite floor would bring considerable economies to the overall cost of the project during its lifetime [10]. The increased popularity of steel framed construction over the last two decades is due to the advantages arising from the use of composite floor. The precast slab panels ideally suited for spans upto 4.2 m, but can be used for large spans by providing secondary beams. For estimation of time and cost of composite floor systems, various authors have been presented the papers [4, 5, 6, 8] in journals.

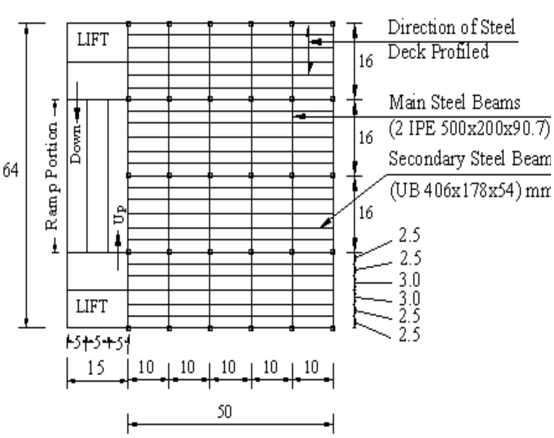
In recent years significant development has taken place in the structural design of multistoried buildings, mainly based on the principles of composite construction. This will improve the speed of construction and reduce the overall construction cost. The main objective of steel framed composite floor construction is to provide a cost-effective alternative to the any other type of construction such as precast slab panel floor.

The building in case study is totally, a 10 storied modern multilevel car parking composite steel structure (Figure 1). It has rectangular (50 m x 64 m) in plan with nominal height 31.5 m (3.15 m floor to floor height) and gross floor area 32000 m2 (3200 m2 at each floor), located at Infosys IT Park, Pune, India. The building is designed and constructed as post-tensioned composite steel beams having 10 and 16 meter spans (Figure 2). The 'Satyam' trapezoidal cold-formed steel deck profile is used for composite floor construction having 11500 mm length, 954 mm width and 1.00 mm thick.

Considering the same plan, total floor area, floor to floor height and loading condition, the existing steel framed composite floor structure is designed and constructed by other two ways;

- Precast concrete frame with precast concrete floor.
- Steel frame with precast concrete floor.

Figure 1: Multilevel Car Parking Steel Framed Composite Floor Building



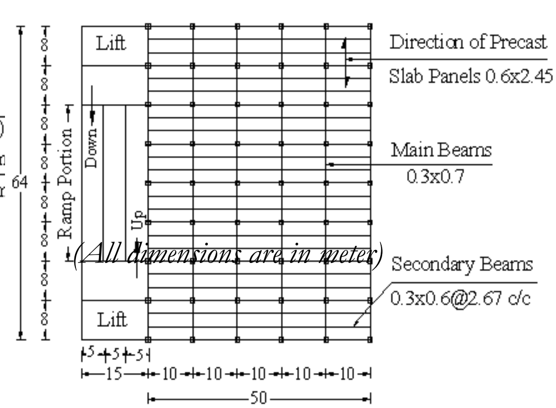


Figure 2: Typical Plan of Multilevel Car Parking Building

Figure 3: Typical Plan for Precast Frame with Precast Concrete Floor

In the first case, the structural members of the building viz. column, beam and slabs are designed and constructed as precast concrete members with Siporex slab panels [13]. For precast building, an additional column is introduced in between two columns of the whole span as shown in Figure 3. In second case, the structural members of the building viz. column and beams are designed and constructed as similar to case study and the construction of slab as a precast concrete floor with Siporex slab panels.

Two significant factors are considered for evaluating composite floor and pre-cast floor building, i.e. optimum time required for the construction and the total cost of buildings. The optimum time has been calculated by using Microsoft Project-2003 [9]. The construction of each structure is divided into various activities, which provide the relative time saving and the optimum time for construction. Considering all above parameters, the material and construction cost of buildings are calculated using market rates in the year 2007, for Pune (India). Structural element wise cost evaluation of all three structures is done as material and construction costs of buildings. However, study of comparison between composite floor and precast concrete floor is restricted to structural frame and slab only.

2. Salient Features of Multilevel Car Parking Building

- It is a totally steel framed structure having capacity up to 2000 vehicles per 10 hours per day for first shift and minimum 1000 vehicles per 10 hours per day for second shift, *viz.* minimum 3000 vehicles per day.
- It is a unique structure in India, as it consist Post-tensioned composite beams. It helps to reduce sizes of the beam sections and also helps to keep larger clear span between two columns.

3. Construction methodology of steel framed composite floor building

Foundation of this building is normal box footings and RCC pedestals to hold steel columns. Anchor bolts of length 1200 mm is provided to hold huge steel columns in a position with special arrangement. Erection of column and beams of this building can be done in four stages. As shown in Figure 4, first lift includes the ground, first and second floor column and beams erection only. While second lift is in progress that time ground, first and second slab construction activities are in a progress. Likewise whole structure can be erected for all four lifts. Time scheduling gives clear idea about such simultaneous activities. After completion of column & beams erection for first lift, at one time construction of composite floor is progresses for three floors. On third floor deck sheet placing, on second floor shear stud welding and on first floor reinforcement lying is in progress. This type of system allows many work faces open together and huge amount of time saving can be achieved. Figure 6 shows the section of composite floor slab with all details.

For pre-stressing of steel main beams, six cables on each side of the beam (includes six tendons in each cable) and for secondary beams, two cables on each side of the beam are placed with the help of fixtures. Post- tensioning operation is done in both directions (50 m x 64 m). In shorter direction cables are tensioned for five spans of 10 m lengths and in longer direction for four spans of 16 m lengths as shown in Figure 5. The construction activity of post-tensioning of composite steel beam is divided into three stages. In first stage, after 14 days of slab casting, the post-tensioning is with only 50% load. In second stage, after 21 days of slab casting, the post-tensioning is with 25% load. In third stage, after 28 days of slab casting, the post-tensioning is with 25% load. The full post-tensioning is done after 28 days of slab casting.

4. Construction Methodology of Precast Frame with Precast Concrete Slab

The structural analysis of precast building were carried by using STAAD Pro-2005 [12] and the precast members are designed as RCC structure [6, 7]. The construction schedule of PCC and RCC footing for precast concrete floor is same as that of steel frame with composite building. Only the numbers of PCC and RCC footings are increased which increased total duration and cost of project. The precast columns used are hollow precast section with sleeves at the top of column section for interlocking of column and beams. Before beams placing the two meter height hollow section of column is grouted with rich concrete of M30 grade by using the self compacted admixture viz. Viscous to reduce the porosity in concrete and above portion is grouted with screed concreting. Rectangular shaped partially precast beams with open stirrups and flanges of 100 mm width are provided for bearing between panel and beam sections (Figure 7). All the precast slab panels are 600 mm wide, 2450 mm long and 125 mm thick. Before placing of reinforcement and screed concrete, a layer of water repellent agent viz. silicon oil is applied on top surface of the panels. The dowel bars for beam, column and reinforcement steel for floor screed is laid on complete floor. The screed of 50 mm thickness is laid on the top of panels with a nominal reinforcement of 8 Φ @ 250 mm c/c having concrete M25 grade.

5. Construction Methodology of Steel Frame with Precast Concrete Floor

All the structural members are designed and constructed according to Eurocode-4 [3], IS 13990 [6] and IS 13994 [7]. As compared with precast frame with precast concrete floor method, shear studs take the function of the dowel bar for beams. The composite action between steel beam and concrete slab through the use of shear connectors is responsible for a considerable increase in the load-bearing capacity and stiffness of the steel beams, which when utilized in design, can result in significant savings in steel weight and construction depth (Figure 8). The headed shear studs [1] as shown in Figure 8, are used with spacing 318 mm c/c. The shear stud welding is done by self-taping machine, which reduce the time of activity; but this method is costly than normal welding method [11]. The screed of 50 mm thickness is placed with the nominal reinforcement of 8 Φ @ 250 mm c/c having M25 grade of concrete. The post-tensioning of steel beam is divided into three lifts, similar to the construction of steel framed composite floor building.

Figure 4: Erected Frame after 1st Lift

Figure 5: Post-tensioning cables at column junction

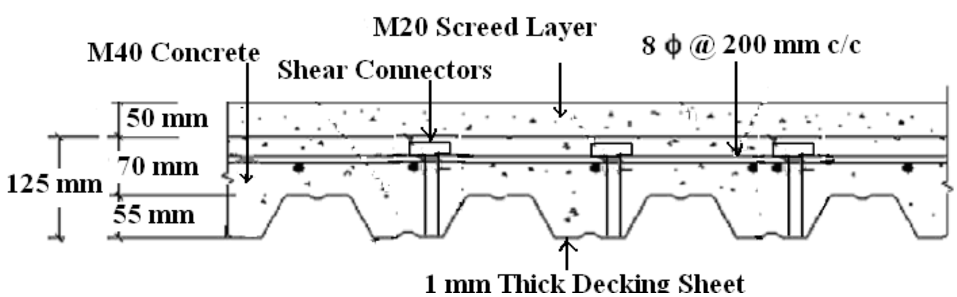
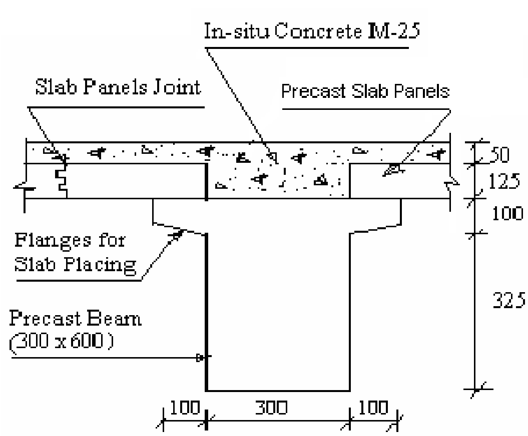


Figure 6: Section of Composite Slab



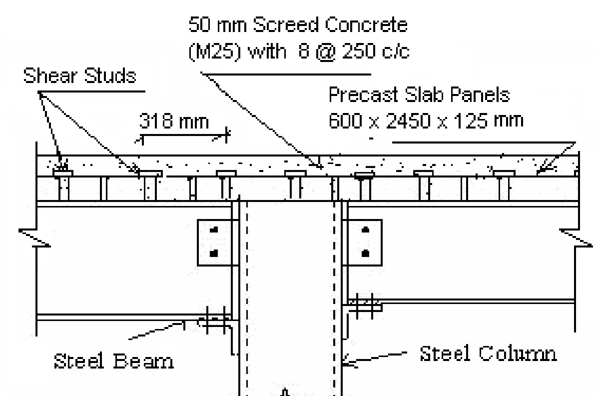


Figure 7: Section of Precast Concrete Figure 8: Section of Steel Frame Beam and Slab Panel with Precast Concrete Floor

6. Time Scheduling:

a) Steel Framed with Composite Floor:

Time scheduling is done using Microsoft Project 2003 [9]. In time scheduling some starting activities such as PCC (7 days), Footing (15 days), and Pedestal (7 days) goes individually but after completing column and beam erections for first lift (33 days), activities for composite floor construction goes simultaneously with second lift erection. Likewise whole structure can be erected with; so many works faces open together.

For the construction of composite floor for all levels, requires 118 days as per time scheduling. It shows that ground and first floor slab activities and for all remaining floors activities are same as first floor. Considering time required for all floors, the building is completed in 180 working days (Figure 9). Total 210 days are required including holidays.

b) Precast Framed with Precast Concrete Floor Building:

Considering the lifting time for precast structural elements, the erection time for column and beams are increased with floor levels, viz. 10 days for first three floors to 12 days for last three floors. Placing of slab panels is started after 2 days of column grouting, it required 6 days for ground floor and 9 days for top floor with same workforce. In time scheduling, 12 days for PCC, 22 days for footings, and 150 days for first six floors construction viz. 25 days per floor and 104 days for next four floor construction viz. 26 days per floor are required. The erection of ground floor column started immediately after 14th day of footing construction, which saves 8 days to total duration. The total project completed in 280 working days (Figure10).

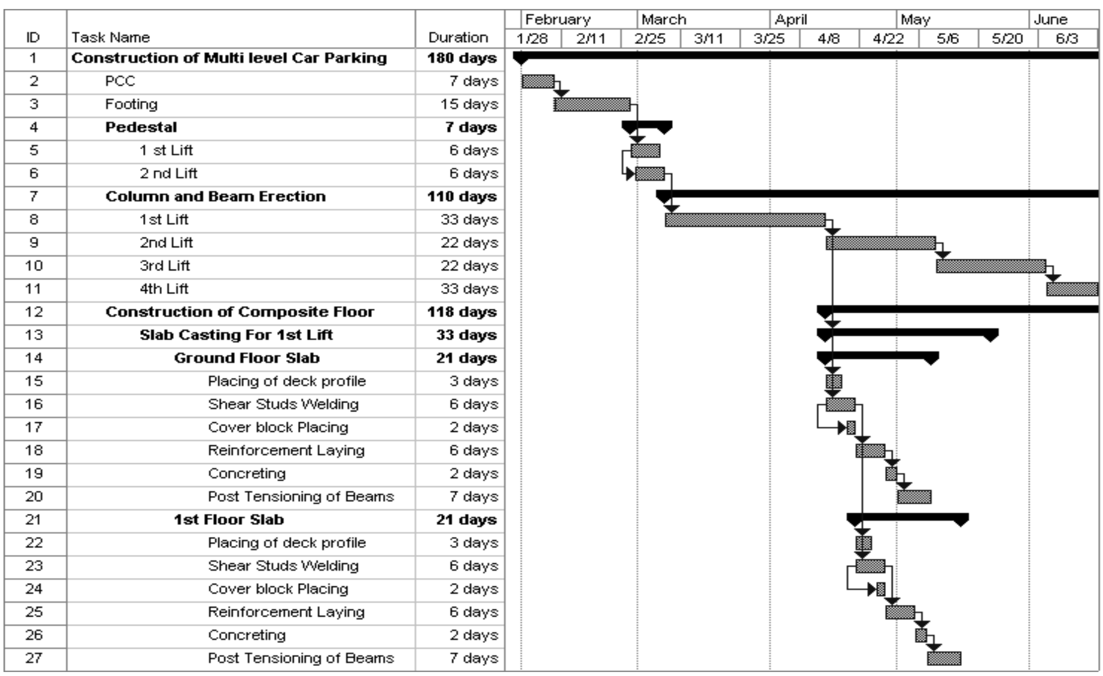


Figure 9: Time Scheduling for Steel Framed Composite Floor Structure

c) Steel Framed with Precast Concrete Floor Structure:

For a partially precast slab construction, 10 workers are required for 10 days for placing precast slab panels, which include lifting, and fixing of slab joints in 14 hrs working time. The concrete for floor is done by using concrete pump with 8 workers including pump operator in 2 days. Post-tensioning of steel beams are planned as same as steel frame with composite floor building. In time scheduling, 7 days for PCC, 15 days for footings, 7 days for pedestal, 110 days for column and beam erections are required and 33 days is delay to start the placing of Siporex slab panel's upto erection of first lift of column and beams, and 143 days for construction of precast concrete floor.

Each slab has slab cycle duration of 32 days but after interlinking of all activities for floor to floor and columns lift to lift, the total duration of three floors become 56 days. The total project completed in 205 working days (Figure11).

As shown in Figure 9, Figure 10 and Figure 11, the time estimation is carried by considering same starting date for all these three type of construction as 1st February 2007, which results into balancing of holidays, working and non-working days. The construction of multilevel car parking building with steel frame composite floor (180 days) saves 55% time than precast frame with precast concrete floor (280 days) and 13% time than steel frame with precast concrete slab (205 days).

ID	Task Name	Duration
1	**Construction of Precast Frame with Precast Floor**	**280 days**
2	PCC	12 days
3	Footing	22 days
4	**Ground Floor Slab**	**25 days**
5	Erection of Column & Beam	10 days
6	Grouting of Column upto 2m Height	4 days
7	Placing of Precast Slab Panels	6 days
8	Grouting of Slab Joint & Silicon oil Apply	4 days
9	Reinforcement Laying & Formwork Erection	6 days
10	Concreting	2 days
11	**1St Floor Slab**	**25 days**
12	Erection of Column & Beam	10 days
13	Grouting of Column upto 2m Height	4 days
14	Placing of Precast Slab Panels	6 days
15	Grouting of Slab Joint & Silicon oil Apply	4 days
16	Reinforcement Laying & Formwork erection	6 days
17	Concreting	2 days

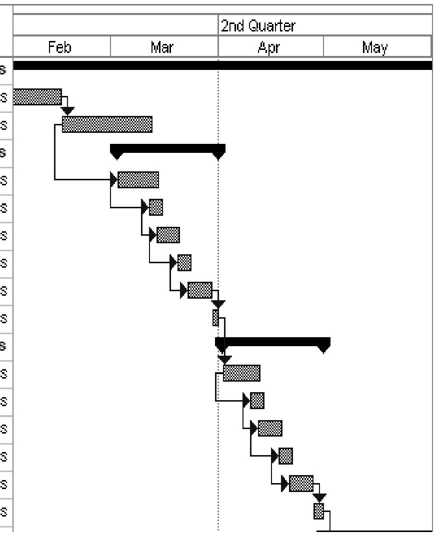


Figure 10: Time scheduling for Precast Frame with Precast Concrete Floor

ID	Task Name	Duration
1	**Construction of Steel Frame with Precast Floor**	**205 days**
2	PCC	7 days
3	Footing	15 days
4	**pedestal**	**7 days**
5	1st Lift	6 days
6	2nd Lift	6 days
7	**Column and Beam Erection**	**110 days**
8	1st lift	33 days
9	2nd lift	22 days
10	3rd lift	22 days
11	4th lift	33 days
12	**Construction of Precast Floor**	**143 days**
13	**Slab Casting For 1st Lift**	**56 days**
14	**Ground Floor Slab**	**32 days**
15	Placing of precast slab panels	10 days
16	Shear Stud Welding	4 days
17	Grouting Slab Joint	4 days
18	Reinforcement Laying & Formwork (	8 days
19	Concreting	2 days
20	Posttensioning	7 days
21	**1st Floor Slab**	**33 days**
22	Placing of precast slab panels	10 days
23	Shear Stud Welding	5 days
24	Grouting Slab Joint	4 days
25	Reinforcement Laying & Formwork (	8 days
26	Concreting	2 days
27	posttensioning	7 days

Figure 11: Time scheduling for Steel Frame with Precast Concrete Floor

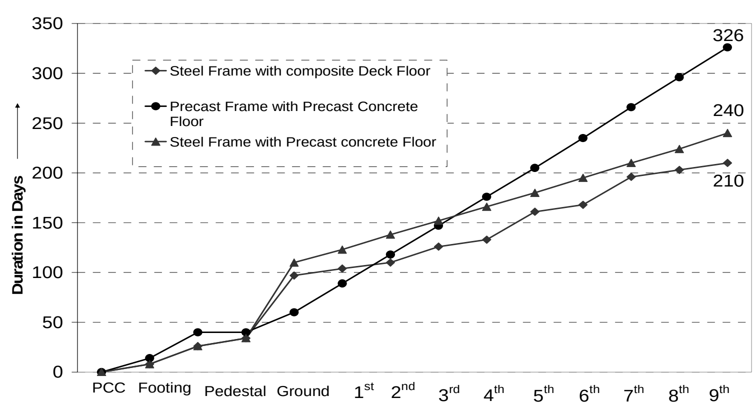


Figure 12: Floor wise Comparison of structures

Figure 12, shows that, after considering one holiday per week, the construction of steel frame composite floor (210 days) saves 55.3% time than that of precast frame with precast concrete floor (326 days) and

14.3% time than that of steel frame with precast concrete slab (240 days).

7. Cost Estimation

a) Direct Cost of Projects:

The total cost of project is divided into four major construction activities such as, foundation, column, beam and slabs construction. Figure 13 and Figure 14 indicate distribution of material and construction cost for these activities, respectively, for each type of building. Same aspect is presented as % of total cost of project in Table 1.

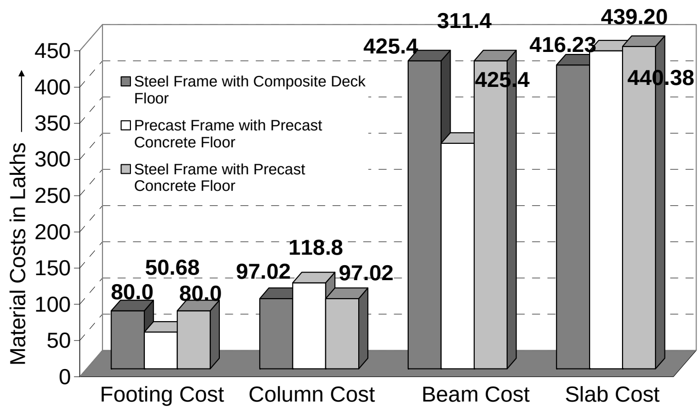


Figure 13: Material Cost of Buildings (As per year 2007)

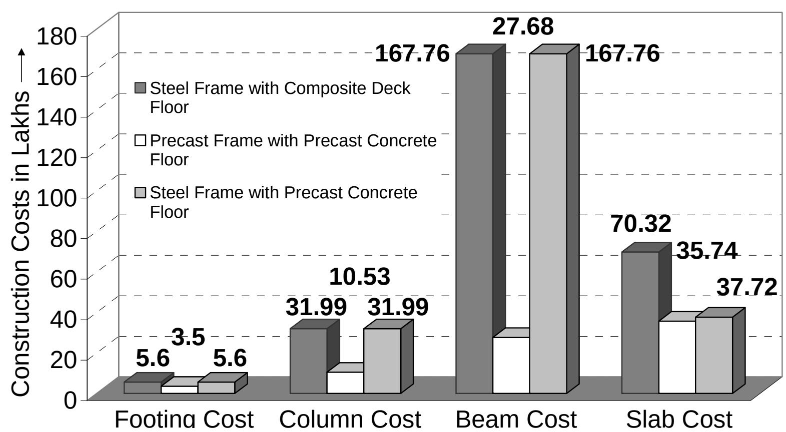


Figure 14: Construction Cost of Buildings (As per year 2007)

Table 1: Percentage of total cost of projects

Structural Elements	Material and Construction cost as % of total cost of projects		
	Steel frame with precast Concrete floor	Precast concrete frame with precast concrete floor	Steel frame with precast Concrete floor
Foundation	7	6	7
Column	10	12	10
Beam	45*	34	46*
Slab	38	48	37

* Steel beams with post tensioning.

It is observed that for the steel frame with composite floor building and steel frame with precast concrete floor structure, the maximum material cost is associated with beams and slabs material while maximum construction cost with erection of post-tensioned beam. For precast frame with precast concrete floor structure, the maximum material cost is associated with slab and beams material and maximum construction cost with erection of slab panels.

Table 2 shows, the percentage increase in material and construction cost of structural elements for other two buildings as compared to steel frame with composite deck floor. The negative values indicate that extra cost is required for it. For precast structure, the total percentage of cost savings are associated with, beam cost (42.83%), footing cost (36.27%), and slab cost (2.38%) and the cost of column is slightly extra by only 0.25% compared to steel frame with composite deck floor. Similarly the percentage of cost savings are associated with beam cost (42.83%), footing cost (36.27%), and a little saving in slab cost as 1.64% as compared to steel framed with precast concrete floor structure.

As compared with steel framed with composite deck floor, the total cost saved through precast frame with precast concrete floor construction is about 23.10%, and the cost saved through steel framed with precast concrete floor construction is only 0.52%. The cost saved for precast frame with precast concrete floor is about 22.70% as compared to steel frame with precast concrete floor. The costs for foundations and costs of columns for all these type of building is least amount as evaluated with cost of beam and slabs construction. The cost of large span post-tensioning composite steel beam affect on the total project cost, project cost increased by 11% of total cost of projects. The reduction in total cost of steel building is attributed to cost of steel being higher than concrete.

Table 2: Percentage Increase in Cost Compared to Steel Frame with Composite Deck Floor

Sr. No.	Structural Elements	Precast Concrete Frame with Precast Concrete Floor (%)			Steel Frame with Precast Concrete Floor (%)		
		Material Cost	Construction Cost	Total Cost	Material Cost	Construction Cost	Total Cost
1	Foundation	36.68	36.88	36.27	0	0	0
2	Column	-22.44	67.08	-0.25	0	0	0
3	Beam + PT Beam	26.80	83.50	42.83	0	0	0
4	Slab	-2.80	48.80	2.38	-6.05	45.99	1.34

b) Net Cost of Projects:

The net cost of project is including the extra cost incurred in interest on barrowed money and car parking rent on saved days of construction, material cost and construction cost of projects, are calculated and compared by considering time related saving, the interest cost on barrowing money and the cost required for car parking rent for saved days of construction. The extra costs of projects are calculated based on interest rate and parking charges in Table 3.

Interest Rate: By taking the survey of various banks and some reputed construction companies, the data for interest rate on the barrowing money is collected and the average interest rate for commercial constructions is considered as 11%.

Parking Charges:

- Parking charges at multilevel car parking building decided by Infosys:
  Rs. 900/- per month per vehicle.
- Parking capacity: 3000 vehicles per day *(as mention in section-3)*
- Extra cost required for car parking rent for saved days,
  *= Days saved in construction * No. of vehicles * parking charges / 30*

The extra cost for all three projects are calculated and compared with the steel frame with composite deck floor. The cost saved for precast frame with precast concrete floor structure is about 12.99% and for steel frame with precast floor structure - 2.32%. The negative value in table indicates extra cost is required than composite deck floor construction. The cost saved for precast frame with precast concrete floor in comparison of steel frame with precast concrete floor structure is about 14.96%

Table 3: Time and Cost of Projects (For the year 2007):

Costs	Steel Frame with Composite Deck Floor	Precast Concrete Frame with Precast Concrete Floor	Steel Frame with Precast Concrete Floor
Total project duration (Days)	210	326	240
Direct cost in Lakhs	1297.37	997.575	1290.679
11% interest on Direct cost in Lakhs	82.120	98.015	93.353
Extra cost required for car parking rent for saved days	0	104.400	27.000
Net cost in Lakhs	1379.484	1199.983	1411.032
Net cost in Rs /m^2	4310	3750	4410
% cost savings as compared to steel frame with composite deck floor		12.99	-2.32
% cost savings as compared to precast frame with precast concrete floor			- 14.96

8. Conclusions

Following are the conclusions drawn out from study:

- The study shows that the time savings of 55.3% is achieved due to use of steel framed composite floor construction rather than precast framed with precast concrete floor and 14.3% time than that of steel framed with precast concrete slab. The construction of steel framed composite floor building saves time, which leads to an overall savings in net cost.
- The direct cost required for steel framed with composite floor is 23.10%, higher than precast frame with precast concrete floor and only 0.52% higher than steel framed with precast concrete floor. Considering time related savings, the net cost required for steel framed with composite floor is 12.99%, more than precast frame with precast concrete floor and 2.32% less than steel frame with precast floor.
- The steel framed with precast concrete floor saves 35.83% construction time than precast frame with precast concrete floor, which required extra 22.70% of direct cost and 14.96% of net cost.
- However, study is restricted to structural frame only. If other items are also considered in the study like excavation work, finishing items, services, cladding etc. and also during construction preliminaries such as labour accommodations, their travelling and food expenses and many other factors related to time, then definitely, steel framed composite floor building option will become cost effective.
- Post-tensioned composite beam is very recent technique and multilevel car parking building might be only structure in India, which has such technique. So cost of material and labour to adopt such new technique is high.
- In present Indian construction sector, there are very less cold-formed trapezoidal profiled steel deck manufacturers. Obviously due to less competition, material rates are much higher. But from present status and already announced investment, future of Indian steel production industry is very bright for cold-formed steel deck sections. So in near future definitely, steel prices will be reduced and steel framed composite floor construction will become competitive in Indian construction sector.

References

[1] Dennis, Lam., Elliott, K.S., and Nethercot, D.A. (2000). "Parametric study on composite steel beams with precast concrete hollow core floor slabs." J. Construction Steel Research, 54, 283-304.
[2] Elliott, K. S., Davies, G., and Bensalem, K. (1993). "Precast floor slab diaphragms without structural screed." Proc., Concrete-2000, Economic and durable construction through excellence, University of Dundee, Scotland, UK.
[3] Eurocode-4 (2003). "Design of composite steel and concrete structures-Part 1.1. General rules and rules for buildings." EN 1994-1-1:2003.
[4] Hicks, S. J., Lawson, R. M., Rackham, J. W., and Fordham, P. (2004). "Comparative Structure Cost of Modern Commercial Buildings." IInd Edition, SCI P137, Steel Construction Institute Publication.UK.
[5] Hedaoo, N. A., Athare, D. S. (2006). "Composite floor system-A cost effective study." J. Concrete India, India Chapter of ACI, 21(4), 3-8.
[6] IS: 13990 (1994). "Indian standard code of precast reinforced concrete planks and joists for flooring and flooring-specification." Bureau of Indian Standard, New Delhi.
[7] IS: 13994 (1994). "Indian standard code of Design and construction of floor and roof with precast reinforced concrete planks and joists." Bureau of Indian Standard, New Delhi.
[8] Langdon Davis (2004). "Multi-storey residential buildings, the cost benefits of steel." Corus Construction Ltd., UK.
[9] Microsoft Project-2003. (2003). "Microsoft Corporation." One Microsoft Way, Redmond, WA98052-7329, USA.
[10] Murunde, Celikag. (2004). "Economic aspects of using steel framed buildings with composite floors: case studies from United Arab Emirates." J. Const. & Building Materials, 18, 383-390.
[11] Report of committee on composite construction. (2002). "Construction considerations for composite steel-and-concrete floor systems." J. Struct. Eng., ASCE, 128(9), 1099-1110.
[12] STADD Pro-2005. (2005). "Research Engineers Ltd.," U.S.A.
[13] Siporex India Pvt. Ltd., Pune, India. "Manufacturers of Concrete Precast Slab Panels."

Development and Application of a Systematic Innovation Procedure for Construction Technology

Wen-der Yu[1], Chih-Ming Wu[2] and Teen-wei Huang[3]

1Professor, Institute of Const. Mgmt., Chung Hua Univ., Taiwan, wenderyu@chu.edu.tw
[2]Ph.D. Student, Institute of Tech. Mgmt., Chung Hua Univ., Taiwan, d09603004@chu.edu.tw
[3]Student, Dept. of CEPM, Chung Hua Univ., Taiwan, b09416044@cc.chu.edu.tw

Abstract

Technology innovation has been recognized as a main driver for the advancement of construction industry. Due to the lack of systematic innovation procedure, the advancement of construction technologies has been slow compared with the other industries, e.g., ICT and Biotech Engineering. This paper proposed a Systematic Technology Innovation Process (STIP) for innovation of construction technologies. The STIP method integrates several techniques adopted for product research and development in other fast innovating industries including patent mapping, root cause analysis, TRIZ, function modeling, simplify design, etc. Details of the proposed STIP method are revisited. The pit-hole repairing technology for road maintenance work is selected for case study. Step-by-step application of the STIP method to the selected construction technology is demonstrated. Deliverables obtained from each step of the STIP is reviewed and evaluated via a technology stage gate (TSG) process, which is commonly adopted in high-tech manufacturing industry. Finally, an innovative design of new pit-hole repairing technology is developed. Evaluation of the innovative technology is also conducted to ensure its feasibility.

Keywords: Technology innovation, patent analysis, TRIZ, process model

1. Introduction

Construction technology was defined as "the combination of construction methods, construction resources, work tasks, and project influences that define the manner of performing a construction operation" [1] to "accomplish a desired aim necessary for human sustenance and comfort" [2]. Robert Harris pointed out that "…there is more to the construction process than just management…there is more to the construction process than just structural design or geotechnical evaluation…[We need] to create better methods for construction…"[3]. Technology innovation can result in revolutionary advancement in construction practice that traditional management techniques and other skills cannot achieve. Therefore, it becomes the critical component for a company's long-term competitive strategy [4].

However, innovation of construction technologies has been slow compared with other areas in Civil Engineering and other industries, e.g., Information and Communication Technology (ICT), Bio Genetic Technology, Nano Materials, etc. (Nam and Tatum, 1989). One of the critical reasons and maybe the most important one is the lack of a systematic approach for fast innovation [5]. As pointed out by Daniel Halpin in his speech of the Seventh Peurifoy Construction Research Award: "…we need a common framework—a common language" [6]. A Systematic Technology Innovation Process (STIP) is proposed to respond the appeals posed by previous researchers. The goal of STIP was to provide a common framework for fast innovation of construction technologies based on modern product innovation methods adopted in other highly innovative industries. In this paper, the STIP method is applied to innovate pit-hole repairing technology for road maintenance work.

The rest of the paper is presented in the following manner: the previous researches on construction technology innovation are reviewed in the second section; the Systematic Technology Innovation Process (STIP) for fast innovation is proposed and described in details; a case study on the application of STIP to innovate a pit-hole repairing technology for road maintenance work is described in the fourth section; finally, conclusions are drawn and future researches are suggested for interested researchers.

2. Construction Technology Innovation

Innovation of construction technologies has resulted in dramatic revolutions in construction practice. For example, the introduction of Portland cement in 1824 has brought up thousands of new construction technologies and equipment that completely change the way of construction engineering; furthermore, in the first quarter of the 20th century, the steel structural technology was invented and introduced to construction industry, which triggered a second revolution of construction technologies. During the late 1970's, construction industry suffered in low productivity, hence inspired the next generation of construction innovation. Issues such as constructability (O'Connor and Miller, 1994), prefabrication, modularization (Tatum et al., 1986), and automation (Sarah, 1997) have drawn numerous researchers to devote in the innovation of construction and management processes.

In spite of tremendous efforts spent, innovation in construction industry has been relatively slow. Lack of a common framework, as pointed out by Halpin, may contribute significantly to this lag. Previous researchers have exploited many approaches for organization process innovation [1], technology evaluation [7], and advanced technology repositories [8]. However, few of these efforts target directly to design of new technologies. Halpin proposed a CYCLONE model for analysis of construction processes [9]. Many efforts on construction process simulation followed him, e.g., COOPS [10] and STROBOSCOPE [11]. Most of the functionalities of process simulation techniques are still limited to the modeling of existing processes, rather than the invention of new technologies.

Just recently, a new area of construction innovation has been developing on patent analysis (PA) [12][13] and the Theory of Innovative Problem Solving (TRIZ) [14][15][16]. The former innovates the target technology based on existing technologies of the other areas, which are stored in public patent databases; the latter applies a systematic procedure to identify engineering potentially improvable attributes with tools provided with TRIZ [17].

Unlike the simulation approach to innovate the existing construction processes, PA- or TRIZ-based technology innovations seek a different dimension of technology improvement. The former belongs to "incremental innovation", and the latter belongs to "system innovation" or "radical innovation" according to the classification of Sarah Slaughter [4]. The "system" or "radical" innovations usually involve tremendous amount of information and knowledge and need to be performed with assistance of computer aided tools [18]. Such tools are incorporated into a systematic technology innovation process called STIP, which will be described in the next section.

3. Proposed Systematic Technology Innovation Process (STIP)

The objective of STIP method is to achieve a fast innovation of construction technologies by integrating three modern techniques: (1) a product research and development procedure called Research and Development Project Management (R&D PM); (2) an inventive problem-solving method namely TRIZ; and (3) a computer aided innovation tool called Goldfire Innovator™. The STIP procedure consists of eight steps described as follows.

3.1 Root Cause Analysis (RCA)

The RCA step analyzes the potential opportunities for improvement with the identified technology problem. This step is associated with the Opportunity Analysis stage of the R&D PM Process. Two CAI tools are employed to perform RCA: the RCA module and knowledge database provided by Goldfire Innovator™.

3.2 Target Technology

The Target Technology step searches the patent database for the root causes determined in the last step. This step is associated with the Concept Definition stage of the R&D PM Process. The patent databases and patent search tools can be employed to identify the target technology.

3.3 Function Modelling

The Function Modelling step constructs the function model (FM) of the target technology identified in the last step. This step is associated with the Conceptual Design stage of the R&D PM Process. The Function Modelling module of Goldfire Innovator™ can be employed to construct the FM of the target technology.

3.4 FM Modification

The FM Modification step modifies the FM of the target technology obtained in the last step. Principles of TRIZ, CT, value engineering, or simplify design can be adopted for this end. This step is associated with the System Analysis and Basic Design stage of the R&D PM Process. The Simplify Design module of Goldfire Innovator™ or any other innovative solution generator (ISG) commercial software can be employed to construct the FM for the target technology. The result of FM Modification is an "innovated alternative" that improve the problem of the target technology.

3.5 Alternative Evaluation

The Alternative Evaluation step evaluates the modified FM of an innovated alternative generated in the last step. The result of evaluation can be "approval" or "rejection". If the alternative is approved, the STIP proceeds to next step—Method Design; on the contrast, should the technology alternative be rejected, the process goes back to FM Modification to generate a new alternative. This step is similar to the technology stage gate (TSG) of the R&D PM Process [19], which provides the innovator a quality control function of product development.

3.6 Method Design

The Method Design step generates feasible solutions for an approved FM of an innovated alternative; that is, suggests a combination of resources (e.g., devices, materials, equipment, and human resources) and process to implement the innovated technology. This step is associated with the Product Design stage of the R&D PM Process. The knowledge database provided by Goldfire Innovator™ can help the innovator in generating technology solutions. Other approaches for Method Design include brain storming, focus group, and expert interviews when the CAI is not available [19].

3.7 Prototyping

The Prototyping step implements the innovated technology generated in the last step with the available resources and methods. The implementation is experimental rather than formal. The objective is to test the feasibility of producing physical and practical methods that can be experimented or tested in the next step. This step is associated with the Prototyping stage of the R&D PM Process.

3.8 Experiment and Testing

The last step of STIP method is Experiment and Testing. In this step, the prototyped technology is tested with real world scenarios to verify its feasibility and applicability. Design of Experiment (DOE) can be adopted to plan the experiments for testing. Modifications and adjustments may be made to the previous steps (Method Design and Prototyping) if the experiment results show potential problems of the prototype technology.

4. Case study

The STIP method was successfully applied to improve a product (or device) type technology, e.g., reinforced concrete (RC) building pipeline leakage repairing technology [5]; however, it has never been applied to innovate a process type construction method. There are two objectives for this case study: first, to investigate the applicability and feasibility of STIP fast innovation method for process type technologies; second, to develop an innovative design for the pit-hole repairing technology for road maintenance work.

4.1 Scope of Case Study

The case study was conducted in Taiwan to innovate the pit-hole maintenance and repairing technology for asphalt concrete (AC) road pavement. Due to the limitation of time, the scope of patent search was limited to USTPO [20].

4.2 Application of STIP Method

(1) Root Cause Analysis (RCA)

The two critical requirements of the existing AC road repairing technology are the confined working zone (to maintain the operation of the road) and limited time available for performing the work. The focus of this study is put on the latter, fast repair requirement, which is identified to be more desirable for practical construction than the former by interviews with the domain experts. An RCA diagram is drawn, as shown in Figure 2, to illustrate the root causes leading to slow repairing works of AC pavement. In Figure 2, there are two roots causing slow road pavement repairing works: (1) inaccurate material supply—this causes adjustment and reapplying required during the repairing work, and those works prolong the repairing time; (2) insufficient strength development of AC material—this cause the extra time required after repairing work is done. Both of the above causes are responsible for the slow repairing work. Noted that the wet conditions in rainy days can also cause delay of road repairs. However, according to the specifications road works in Taiwan, it is not allowed to perform road pavement work under rains. As a result, the wet condition is not considered in the RCA.

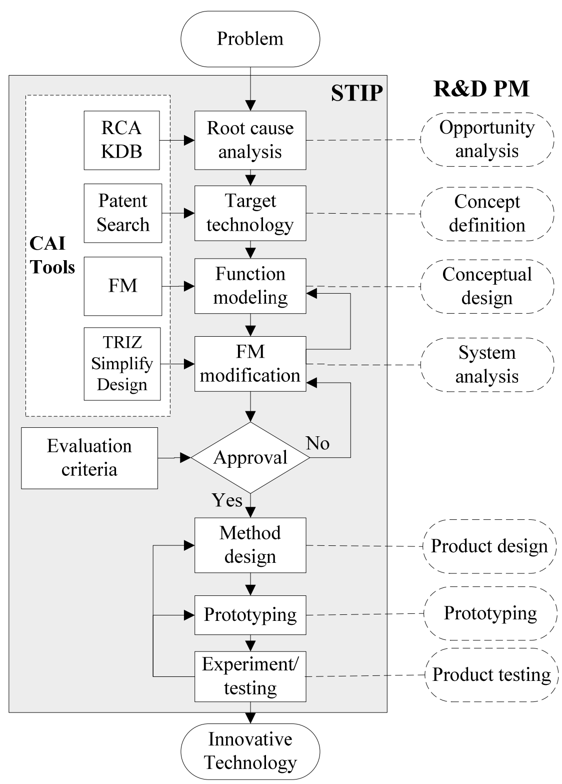


Figure 1 STIP procedure

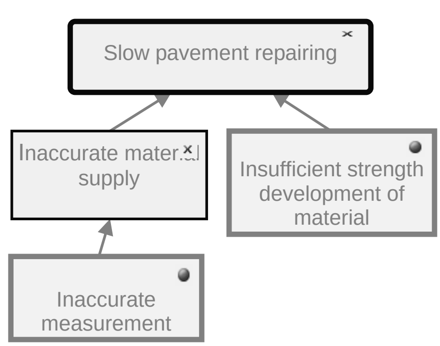


Figure 2 RCA diagram for slow road repair works

Further investigating the cause of "inaccurate material supply", it is found the inaccurate measurement of the material required is the root. Therefore, providing an "accurate measuring method" is the key to solve the "inaccurate material supply" problem. Similarly, providing a high early-strength AC material can solve the "insufficient strength development of material" problem.

The root causes identified by RCA method is used in the next step to search for candidate technologies.

(2) Target Technology and Patent Search

In this step, the patent databases are searched to find out the most appropriate technology that can be considered as the “target technology” for innovation. At first, patent maps are developed to visualize the status of the technological competitiveness in the considered technology domain.

Table 1 shows the search criteria for pavement repairing technologies used in this research to find out the relevant patents in USPTO. The International Patent Classification code (IPC) was adopted in the search. In Table 1, it is noticed that the IPC class: “E01C 23/00” (build, repair, fix, rehabilitate, or demolition of road or similar facilities) was found to be most relevant to the problem domain of the case study.

Table 1 Search criteria for pavement repairing technologies

(((TTL/repair OR ABST/repair OR ACLM/repair OR TTL/rehabilitate OR ABST/rehabilitate OR ACLM/rehabilitate OR TTL/mend OR ABST/mend OR ACLM/mend OR TTL/renew OR ABST/renew OR ACLM/renew)) OR ((ABST/pavement OR TTL/pavement OR ACLM/pavement OR TTL/way OR ABST/way OR ACLM/way OR TTL/road OR ABST/road OR ACLM/road))) AND (ICL/E01C23/00 OR ICL/E01C23/06):579 patents
E01C 23/00: build, repair, fix, rehablitate, or demolition of road or similar facilities.

The search results are used to constructed patent maps (Yu et al., 2006) so that the competitors of the technology domain can be identified. Some of the patent maps are shown in Figure 3 to 5. The analysis results showed that the patent activity chart shows that road repairing technology is declining in the past five years and top three competitors in the technology domain are Eigenmann, Wirtgen, and CMI.

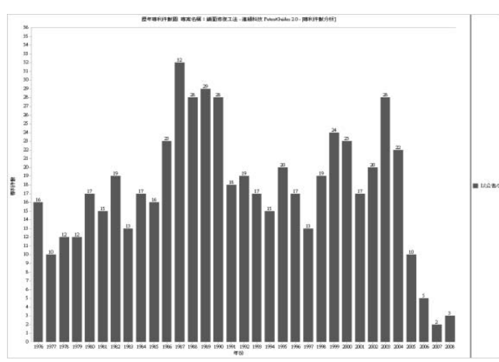


Figure 3 Patent quantity comparison chart (Publication date)

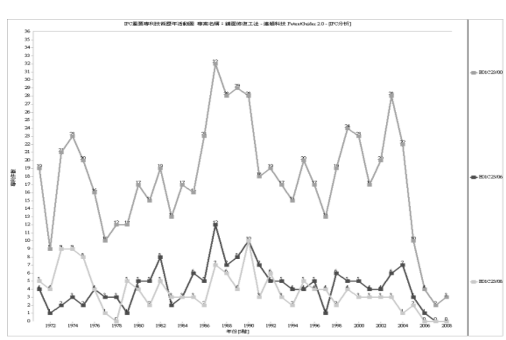


Figure 4 IPC patent activities (TOP 3)

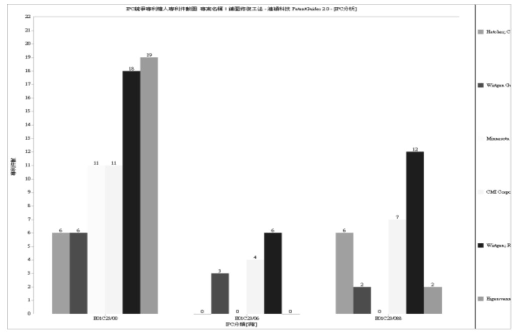


Figure 5 IPC patents of competitor companies (IPC TOP3, Asignee TOP 6)

By reviewing the most relevant patents, the "US4084915: Method for reconditioning and resurfacing pavement" is selected as the target technology for innovation. Since the patent documentation published by USTPO does not describe the construction process of the target technology, domain experts are consulted and a Hot-mixed AC Refilling Method for AC pavement repairing is conceived to be most relevant to the target technology. The construction process is described as follows.

Material Requirements: refilling material should be hot-mixed dense graded AC with aggregate of maximum diameter 13 mm.

Construction method:

- Clean deteriorate material of the damaged portion and the surrounding area of pit-hole with mechanical cutter. The cutting face should be plane.
- Remove the loose aggregate and sundries.
- Coating the cutting surface with a cohesive layer (the cement asphalt mortar can be used).
- Fill in hot-mixed AC material to the pit-hole. Flat the material to out stand the repair surface for 2~3 mm.
- Compact the repair surface with roller.
- Curing until the development strength of the material is sufficient for operation.

(3) Function Modelling

The target technology is converted into function model for further analysis. Since the target technology is more relevant to a procedural method rather than a equipment or device, the Process Model (PM) provided by Goldfire Innovator™ is adopted for function modeling and technology representation of the target technology. Figure 6 shows the PM of the target technology based on the construction method described in the last step. Notice that Provide Link (Prv) implies that the preceding process provides inputs for the successor process; while Corrective Link (Co) implies that the preceding process corrects (or modifies) the functions for the successor process.

In Figure 6, the first step (Clean deteriorate material of pit-hole) provides working space for the second and the third steps; and the last step (Roller compacting) corrects the work results of the fourth step (Fill in hot-mixed AC material).

(4) FM Modification

It was identified by RCA that the root cause for "inaccurate material supply" is "inaccurate measurement". This problem happens at the fourth step of the PM in Figure 6. The computer aided innovation tool (with Goldfire Innovator™) suggests that a new alternative can be developed to substitute the original method. Applying the contradiction matrix of TRIZ, it is obtained that "Improving EP-28 (measurement accuracy)" results in "Deteriorating EP-25 (waste of time)". The suggested inventive principles (IPs) are "IP-24: Mediator", "IP-34: Rejecting and Regenerating", "IP-28: Replacement of Mechanical Syste", and "IP-32: Changing the Color". Considering the above principles, a new operation is adopted to replace the fourth step in Figure 6. The resulted modified process model is shown in Figure 7. Notice that the "Fill in hot-mixed AC material" of the original process is replaced with a new operation. The new operation of the fourth step adopts a laser scanner and associated software to measure the volume of required fill-in AC material; the material is supplied with an automated equipment. Such conceptual alternative will be realized in Method Design.

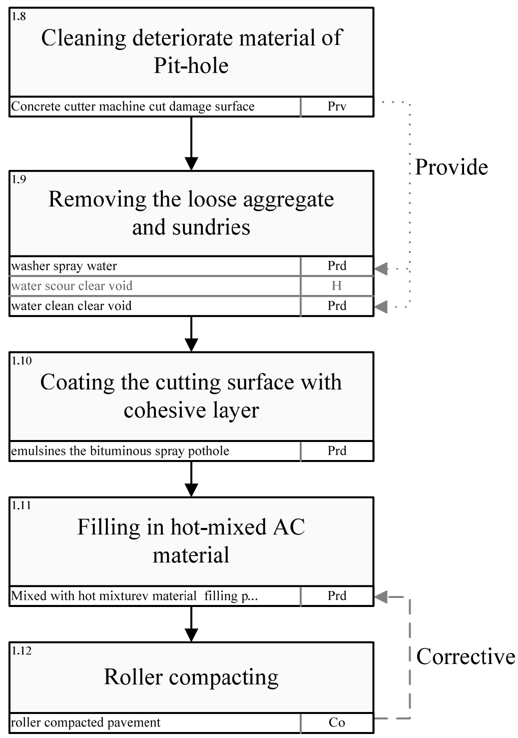


Figure 6 Process Model of the target technology

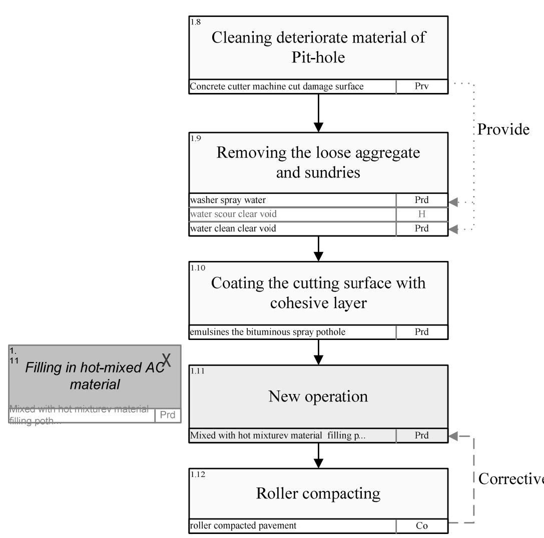
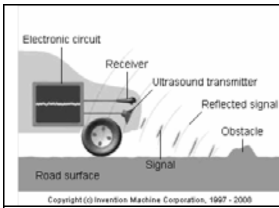


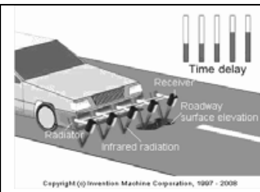
Figure 7 Modified PM

The idea of new operation can be generated by requesting the Knowledge Database with "How to fill the pothole?", the suggested solutions are shown in Figure 8. In Figure 8, four solutions are suggested: (1) Obstacle size affects reflected wave intensity—a wave intensity sensor can be employed to detect the obstacle; (2) Infrared radiation detects roadway surface elevations—Infrared radiation device can be employed; (3) Reflected light detects road irregularities—light detector can be employed; and (4) Laser pumping device—suggesting that laser device is applicable. The fourth solution was obtained by tracing back the "Effect Chain" of the Science Effect database, which illustrates the underlying principle of the first three solutions.

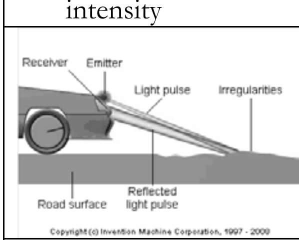
Similarly, four published patents were suggested by Goldfire Innovator™: (1) US 5294210-Automated pothole sensing and filling apparatus; (2) US 5439313-Spray patching pavement repair system; (3) US 6821052 B2-Modular, robotic road repair machine; and (4) US 4511284-Pothole patcher. These are shown in Figure 9.



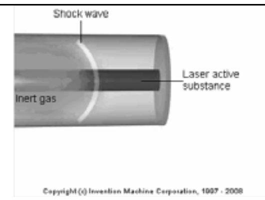
(1) Obstacle size affects reflected wave intensity



(2) Infrared radiation detects roadway surface elevations

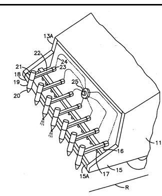


(3) Reflected light detects road irregularities

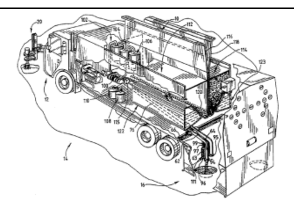


(4) Laser pumping device

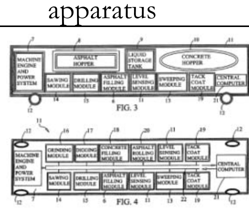
Figure 8 Solution suggested by Sciecne Effect database



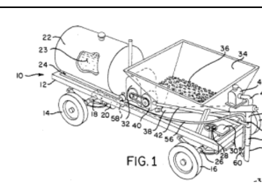
(1) US 5294210 Automated pothole sensing and filling apparatus



US 5439313 Spray patching pavement repair system



US 6821052 B2 Modular, robotic road repair machine



US 4511284 Pothole patcher

Figure 9 Solutions suggested by previous patents

(5) Alternative Evaluation

The alternative evaluation is performed qualitatively with the domain experts in terms of functionality, constructability, and cost effectiveness. The evaluation results are shown in Table 2. The measurement of required material volume and material supply of the original technology was performed manually by the laborers. They are replaced by automatic equipment and technology. As a result, the functionality and constructability are improved. However, the new operation requires additional equipment, which will increase the cost, and thus the inferior cost effectiveness.

Table 2 Evaluation of the innovated technology

Criterion	Technology	
	Original	Innovated
Functionality	Medium	Good
Constructability	Medium	Good
Cost effectiveness	Good	poor

(6) Method Design

In this step, the implementation method for the conceptual innovation technology is designed. The Computer Aided Innovation (CAI) tool, Goldfire Innovator™, is counseled again to generate design scenarios. From Figure 8 and 9, the IP-28 of TRIZ suggests that a laser scanner can be employed for measurement of the pit-hole volume and required material; similarly, the IP-24 suggests that computer software can serves as mediator that can improve the accuracy of measurement and supply of required material. Both of the two functions are available in the Science Effect and patent databases. However, there has been no design to combine the two functions in road repairing. A conceptual design of the innovated technology is shown in Figure 10.

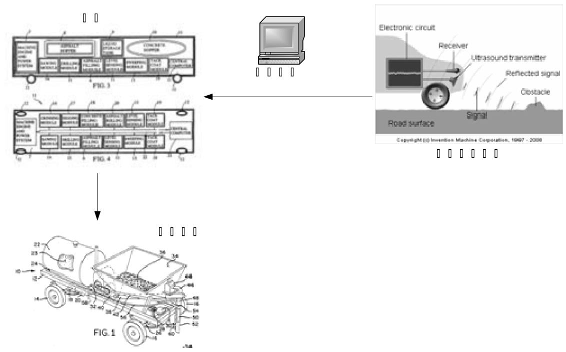


Figure 10 Method design of the innovated technology

(7) Prototyping

The innovated technology has not been physically implemented yet, but a prototype illustration of the innovated technology is shown in Figure 11. The prototype technology consists of four major components: 1) a laser scanner—that detects and scans the pit-hole; 2) a AC material remover—that cleans and removes the deteriorate material of pit-hole scans the pit-hole; 3) a computer with required software—that calculates the volume of pit-hole using the scanned data; and 4) an AC patcher—that fills in AC materials with the required volume and compacts the surface.

(8) Experiment and Testing

Until the deadline of paper submission, the prototype technology was not experimented and tested yet. It will be part of future work. A patent application for the innovated technology is filed to the Taiwan Intellectual Property Office (TIPO) after conceptual design is finished.

5. Conclusions

In the paper a proposed STIP method for fast innovation of construction technologies is described in details. Unlike the traditional simulation-based technology improvement techniques, the STIP method generates alternative technologies based on inventive problem-solving techniques (e.g., TRIZ) and technology databases (e.g., Science Effect and patent databases). As a result, it achieves the “radical” or “system” innovation of construction technologies as classified by Slaughter [4].

A process innovation case study of STIP to innovate the pit-hole repairing technology of AC road pavement is conducted to verify and test the proposed STIP. By following the STIP procedure, an innovative alternative for the target technology is successfully generated and designed. It is concluded that the proposed STIP method is feasible and applicability for innovation of process type technology such as road pavement repairing. It is convinced that such method can also be employed to innovate other types of construction processes and technologies.

Although the conceptual prototype of the innovated technology has been developed, real world implementation and experiment should be conducted to verify the proposed prototype. This will become the future works. Moreover, evaluation of the innovated technology was performed qualitatively in this paper, quantitative evaluation will be performed with the technology experiment and testing in future works, too.

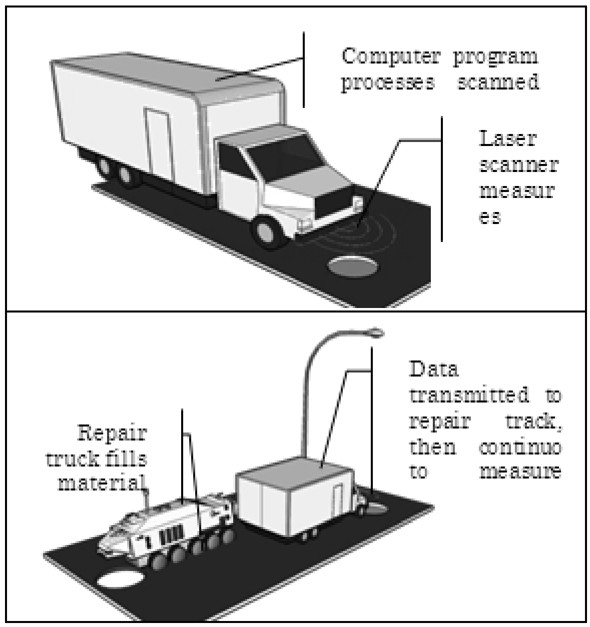


Figure 11 Illustration of the innovated technology

References

[1] Tatum, C. B. (1987). "Process of innovation in construction firm," J. Constr. Engrg. and Mgmt., ASCE, Vol. 113, No. 4, 648-663.

[2] Shin, H., Watanabe, H., and Kunishima, M. (1989). "A new methodology for evaluating a new construction technology from the viewpoint of constructability," Proceedings of the 47th Doboku Gakkai Rombun-Hokokushu, Japan Society of Civil Engineers April 1989, JSCE, Japan, 4 pp.

[3] Harris, R. B. (1992). "A Challenge for Research," J. Constr. Engrg. and Mgmt., ASCE, Vol. 118, No. 3, 422-434.

[4] Slaughter, E.S. (2000). "Implementation of construction innovations," Building Research and Information, Vol.28, No.1, p2-17.

[5] Yu, W. D., Wu, C. M., and Lien, W. C. (2008). "Fast innovation of construction technologies with computer aided innovation tools," Proceedings of International Symposium on Automation and Robotics in Construction 2008 (ISARC 2008), June 27~29, 2008, Vilnius, Lithuania, pp. 521~527.

[6] Halpin, D. W. (1993). "Process-based Research to Meet the International Challenge," J. Constr. Engrg. and Mgmt., ASCE, Vol. 119, No. 3, 417-425.

[7] Yu, W. D., and Skibniewski, M. J. (1999). "Quantitative constructability analysis with a neuro-fuzzy knowledge-based multi-criterion decision support system," Automation in Construction, 8(5), 553-565.

[8] Ioannou, P. G. and Liu, L. Y. (1992). "Advanced Construction Technology System—ACTS," J. Constr. Engrg. and Mgmt., ASCE, Vol. 119, No. 2, 288-306.

[9] Halpin, D. W. (1977). "CYCLONE: method for modelling of job site processes," J. Constr. Div., ASCE, Vol. 103, No. 3, 489-499.

[10] Liu, L. Y. (1995). "Simulating construction operations of precast-concrete parking structures," Proceedings of the 1995 Winter Simulation Conference, 1004-1008.

[11] Martinez, J. C. and Ioannou, P. G. (1994). "General purpose simulation with STROBOSCOPE," Proceedings of the 1995 Winter Simulation Conference, 1307-1313.

[12] Rothe, C. A. (2006). "Using Patents to Advance the Civil Engineering Profession," Civil Engineering, ASCE, June 2006, 66-73.

[13] Yu, W. D., Cheng, S. T., Shie, Y. L., and Lo, S. S. (2006). "Benchmarking Technological Competitiveness of Precast Construction through Patent Map Analysis," Proceedings of International Symposium on Automation and Robotics in Construction 2006 (ISARC 2006), Session B2—Design, Planning and Management System, Oct. 3~5, 2006, Tokyo, Japan, pp. 119~123.

[14] Cheng, S. T., Yu, W. D., Wu, C. M., Chiu, R. S. (2006). "Analysis of Construction Inventive Patents Based on TRIZ," Proceedings of International Symposium on Automation and Robotics in Construction 2006 (ISARC 2006), Session B3—Construction Planning and Management Methods, Oct. 3~5, 2006, Tokyo, Japan, pp. 134~139.

[15] Mohamed, Y. and AbouRizk, S. (2005). "Technical Knowledge Consolidation using Theory of Inventive Problem Solving," J. Constr. Engrg. and Mgmt., ASCE, Vol. 131, No. 9, 993-1001.

[16] Mohamed, Y. and AbouRizk, S. (2005). "Application of the Theory of Inventive Problem Solving in Tunnel Construction," J. Constr. Engrg. and Mgmt., ASCE, Vol. 131, No. 10, 1099~1108.

[17] Altshuller, G. (2002). 40 Principles: TRIZ Keys to Technical Innovation, Technical Innovation Center, Boston, MA, USA.

[18] Li, Y., Wang, J., Li, X. L., and Zhao, W. (2007). "Design creativity in product innovation," International Journal of Advance Manufacture Technology, Vol. 33, p213-222.

[19] TPMA (2007) International R&D Project Management Body of Knowledge, Taiwan Project Management Association, Kaohsiung, Taiwan. (in Chinese)

[20] USPTO, (2008). Web Site, http://www.uspto.gov/, United States Patent & Trademark Office, visited 2008/3.

A Spatiotemporal Approach to Managing Utility Work Schedules

Chien-Cheng Chou[1], Yi-Ping Chen[2] and Chien-Ming Chiu[2]

[1]Assistant Professor, Department of Civil Engineering, National Central University, Taiwan; PH: 886-34227151 ext.34132; FAX: 886-34252960; e-mail: ccchou@cc.ncu.edu.tw
[2]Graduate Research Assistant, Department of Civil Engineering, National Central University, Taiwan

Abstract

As more and more utility installation and/or maintenance activities are located in highly congested urban roadways, frequent pavement utility cuts in such areas may cause more traffic disruption as well as deteriorate pavement life and quality. Utility owners normally need to obtain permits from public road authorities before commencing utility activities; however, public road authorities in Taiwan currently just issue permits without trying to coordinate and communicate with utility owners involved to schedule their utility-related activities in a more consecutive way. An information model based on the spatiotemporal objects database technique was proposed to help public road authorities identify the utility activities that might be combined together to avoid unnecessary pavement utility cuts. In the proposed model, constraints pertaining to pavement moratorium, utility clearance distance and traffic conditions were considered. The software architecture is discussed, followed by research conclusions.

Keywords: Pavement, Utility cut, Utility permit, Spatiotemporal objects database, Temporal database

1. Introduction

As more and more people dwell in urban areas, there is an increasing number of utility installation and/or maintenance activities located in such areas that make a great impact on paved roads. The steady escalation of the internet penetration rate demonstrates the need that more communication equipment such as fiber broadband lines or wireless access points will be deployed along major transportation systems in the near future. In order to provide new services or maintain deteriorating utility networks, utility owners have to cut pavements open, install new utility facilities or fix problems identified, backfill proper materials, and restore road surfaces. Reports showed that in the District of Columbia, there were over 5,000 utility cuts in 1996 and over 6000 cuts in 2000 (Wilde et al. 2003); in New York City, more than 250,000 cuts a year were made in 2000, and the number increased by 8% each year (Khogali et al. 1999). Researchers indicated that utility activities in the U.K. rank as the second major cause of traffic disruptions with estimated delay costs of $13 billion dollars. Additionally, uneven pavement surfaces due to frequent utility cuts may further result in driver annoyance and other safety issues (Jensen et al. 2005). Others pointed out pavement utility cuts as a major problem in the transportation infrastructure of the U.S., creating serious financial stress on public road authorities (Wilde et al. 2003). In the M-Taiwan project, the government in Taiwan planned to build 6,000 km of optical fiber transmission lines during 2004-2007 to provide most residents with high-speed internet services (Lin 2005). All of the above studies reveal that pavement utility cuts are inevitable and may bring about numerous challenges associated with costs, safety, pavement maintenance, etc.

Since utility owners normally need to obtain permits from public road authorities before commencing utility activities, encouraging utility owners to work together is generally recognized by public road authorities as a potential solution to reduce unnecessary pavement utility cuts. However, in Taiwan, most public road authorities currently just issue permits to utility owners without trying to coordinate and communicate with utility owners involved to schedule their utility-related activities in a more consecutive way. Hence, a managerial tool that can keep track of the schedule and geometric boundary of every planned utility activity is highly desired. Any potential cooperation between utility owners that might reduce unnecessary pavement utility cuts would be detected by such tool, and public road authorities can use the suggestions generated to persuade the utility owners into working jointly.

To this end, this research aims at investigating an information model that can help public road authorities manage utility activities to reduce unnecessary pavement utility cuts. Literature review regarding problems associated with utility activities and their overall impact is described first. The proposed information model that is designed to best describe spatial and temporal properties of planned utility activities is presented next, followed by model exploitation and evaluation of the prototype's software architecture. The tasks required to validate the model are described, and research conclusions are made finally.

2. Literature Review

Several approaches to reducing unnecessary pavement utility cuts have been proposed and investigated. All were designed to cause minimum disturbance to traffic and to have a low impact on the environment. Briefly, these approaches can be categorized into two types: technology-based and policy-based approaches (Wilde et al. 2003). The technology type of approaches such as the trenchless technique focuses on construction methods, practices, tools, etc. that can be employed to perform the utility work. The policy type of approaches focuses on how to allocate construction resources, how to manage organizations and teams, and how to communicate and coordinate with project stakeholders in order to control frequent pavement utility cuts. The technology-based approaches usually accompany high initial cost and short history of proven success, whereas the policy-based approaches often involve incentives, disincentive, and changes of permit procedures that utility owners must follow to complete their works (Wilde et al. 2003). In fact, more researchers investigate the technology-based approaches. The policy-based approaches have less attention in the literature, and few researchers recognize the trend that since the number of pavement utility cuts is escalating, public road authorities might need a managerial tool that can assist project managers in condensing or rescheduling the work schedules of planned pavement utility cuts performed by different utility owners in order to minimize the impact on the traveling public. For example, assume that utility company A will install new facilities in a street during certain days, and utility company B will perform maintenance activities in the same area but one month later after completion of A's work. If both A's and B's work schedules are flexible, the public road authority might be able to persuade A or B to reschedule their work to perform consecutively. Identifying the project circumstances where two or more different utility activities can be combined together is very important to public road authorities because the influential area of pavements due to utility cuts can be carefully calculated so that both the number of utility cuts and the affected area can be minimized.

In addition to reducing unnecessary pavement utility cuts, public road authorities also face a challenge regarding the increasing number of interferences among the planned but not-yet-deployed utility facilities. Public road authorities or other public agencies may prescribe the clearance distances between certain types of utility lines. For instance, the Taipei City government regulates any gas pipeline to have at least 5-6m of horizontal separation. To calculate the influential area of a pavement utility cut requires consideration of these planned utility facilities. Research showed that utility permit procedures may take significant time because public road authorities need to consider myriad factors determining whether to issue the permit and coordinate with other organizations to address concerns such as environmental and archaeological issues (Chou et al. 2007). Assume that a utility company is in the progress of acquiring its permit for placing new pipelines, and another utility company would like to submit its permit application to install new facilities. Without proper coordination and communication within the two utility owners, the public road authority might not be able to detect any possible clearance violation of the new utility facilities and still issue the permits. Hence, additional pavement utility cuts may be needed when one of the utility owners performs adjustments to fix the problem. The clearance violation might be resolved if the road segment involving the two utility owners will be rehabilitated jointly because there will be a coordination meeting hosted by the public road authority to address each party's concerns. This is due to the fact that sometimes utility owners are willing to discuss with each other if the public road authority is involved. Intermediate utility facilities to keep up the service during utility work are another possible source of utility interferences. For instance, if water main lines are underneath temporal power distribution poles, serious problems such as voltage shortage or overloading may happen. Therefore, managing the interferences of planned utility activities is becoming a total nightmare from public road authorities' perspective. To reduce unnecessary pavement utility cuts, public road authorities require a systematical approach to effectively and efficiently manage any future utility installation and/or maintenance activities.

Overall, if public road authorities would like to better manage the utility activities, they might need to deal more with the utility work schedules and use a computerized tool to precisely depict the spatial information of each planned utility activity. The following section elaborates more temporal and spatial requirements of each utility activity so that an information model to capture such requirements can be proposed.

3. Temporal Properties of a Utility Activity

To schedule a set of activities, traditional techniques such as Program Evaluation and Review Technique (PERT) emphasize the importance of the constraints on the activities. Constraints involve time or resources-related rules that govern the execution sequences to complete the project. From public road authorities' perspective, since utility activities are performed by different utility owners, sharing resources among these companies is rare, and oftentimes utility owners do not want to overlap their work schedules. The primary constraints in this type of work thus become spatiotemporal-related requirements. For instance, if a utility company will install a new service line in the area, is there any active construction plan located in the same place? When will the plan start and when will it end? The location and the duration are the major spatiotemporal resources that cannot be easily shared in most of the utility activities. If one activity needs a particular spatiotemporal resource, the other activities cannot occupy without proper coordination.

One advanced database technique that has emerged as a main focus of many spatiotemporal information systems such as the digital battlefield in the military is to keep track of object locations over time and to support temporal queries about future locations of the objects (Wolfson et al. 1998)(Wolfson 2002). Called moving objects database (MOD) or spatiotemporal objects database (SOD), this technique aims to deal with geometries changing over time and to simplify the data update process through use of dynamic attributes (Guting and Schneider 2005), thereby having the potential for eliminating or reducing some of the associated challenges and complications. The way SOD employs to process the time dimension for each moving object may serve as a starting point for utility activity modeling.

In SOD, there are two time dimensions associated with each time-sensitive attribute: valid time and transaction time (Guting and Schneider 2005)(Tansel 1993). The valid time refers to the time in the real world when an event occurs or a fact is valid. The transaction time refers to the time when a change is recorded in the database. Formal definitions for Dv (valid time) and Dt (transaction time) are listed as follows:

Dt = {t0, t1, …, ti, …, now}

$$\forall t', t'' \in D_t \setminus \{now\} : t' < t'' < now \vee t'' < t' < now \vee t' = t'' < now$$

Dv = {t0, t1, …, ti, …, now, …}U{∞}

$$\forall t', t'' \in D_v \setminus \{\infty\} : t' < t'' < \infty \vee t'' < t' < \infty \vee t' = t'' < \infty$$

For example, the duration of a utility activity is defined by a manager as from January 1, 2009 to March 1, 2009, which is in the Dv domain. The manager enters this activity information, including the geometric boundary and the schedule, to a computer on November 1, 2008, which is in the Dt domain, but the schedule is changed on January 5, 2009. The new schedule is from February 1, 2009 to April 1, 2009. There is another utility activity scheduled to be performed from March 1, 2009 to May 1, 2009. The manager enters the second utility activity to the computer on December 1, 2008 (see Figure 1). In fact, the three database records regarding the two utility activities must be persisted individually since each data entry may be associated with a permit application and fees. A public road authority may ask the utility owner of the first activity to pay additional fees for the time period between January 1, 2009 and January 5, 2009, although no one indeed performs any work at that time. Recording the two time dimensions of each utility activity helps project stakeholders retrieve not only latest date but historical one for future auditing purposes.

Further, the duration of a utility activity may be associated with a different time scale. For example, if an accident destroys a power switch facility, the problem due to switching over voltages needs to be fixed within several hours. However, leaks of a water main lines may take months to fix the problem. The need of a multi-scale time hierarchy associated with each utility activity is evident; however, current relational database techniques cannot provide an appropriate solution to satisfy this need.

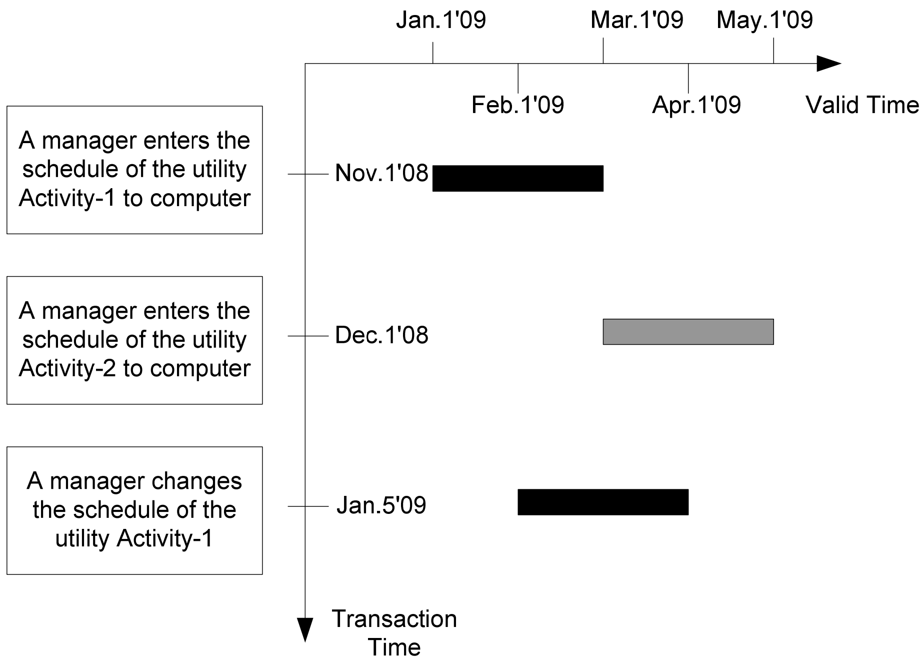


Figure 1. Valid time and transaction time for two utility activities

4. Software Architecture Analysis

In addition to the temporal requirements of a utility activity usually not implemented in current GIS commercial tools, traditional GIS techniques do not fit the internet architecture due to lack of support of the multi-user environment. Since utility permits are often created by respective utility owners at their offices, the modern three-tiered architecture aiming at dividing information presentation, processing and database operations into three layers of software components might be more adequate for our research problem. Based on the three-tiered architecture, a spatial database that can provide both common relational database and GIS functions is needed. An open-source spatial database, i.e., PostGIS, is employed for this research so that the research team can customize some default functions. Briefly, PostGIS is an add-on component designed to conform to the Open Geospatial Consortium (OGC) specification. Its relational database functions are provided by PostgreSQL, a prestigious open-source relational database. Using PostgreSQL with PostGIS, one can easily reuse the robust relational and spatial database engines. As noted before, traditional GIS techniques concentrate on providing an integrated solution to users, include map data presentation, processing and persistence; however, prevalent internet applications change this traditional approach of GIS. Indeed, the new OGC specification covers almost all GIS functions pertaining to data processing. Commercial database products such as Oracle also apply the same architecture as PostgreSQL and PostGIS.

5. Proposed Model

The proposed information model is shown in Figure 2. Elaboration of each class in the model is described in this section. Basically, the model consists of three main parts: (1) roadway classes, including RdNetwork, Road, RdSegment, MaintenancePlan, and Resurface; (2) utility classes, including UtilPart, UtilLine, UtilFacility, Abandonment, UtilSection, UtilNetwork, UtilEnterprise, and UtilDivision; and (3) permit classes, including Permit and UtilActivity. The roadway classes contain information regarding a road network itself and its maintenance plans. The utility classes are designed to capture existing utility lines and facilities, including abandoned ones that are still buried in the field. The permit classes are the reflection of the future utility plan that may have conflicts with other permits, existing utilities, and road maintenance plans. The elements of the model are described as follows:

RdNetwork: This class represents the concept of a road network. A road network object is designed to group road objects. For instance, in a city, the public road authority has four divisions. Each division is responsible for managing one quarter of the city's roads. Hence, four road network objects exist in the model.

Road: This class represents the concept of a road, e.g., Riverside Drive. A road object contains a set of road segment objects. The road class has six attributes. The "Type" attribute is designed to record the road's type, e.g., way, line, drive, road, street, avenue, and boulevard. The "Line" attribute is a "linestring" type in PostGIS, which describes a road as a set of lines. The last three attributes are used to record the traffic flow information of a road.

RdSegment: This class represents the concept of a road segment. A road segment object is the building block to construct a road object. The road segment class has four attributes. The "Lane" attribute is designed to record how many lanes the given road segment has. The "Polygon" attribute can be used to depict the geometry of the segment. The "ROWPolygon" attribute can be used to depict the right-of-way (ROW) geometry of the segment. The "BeginDate" attribute is used to record when the road segment is open to serving the traveling public. Note that this attribute pertains to the valid time dimension so the "EndDate" attribute is set to forever.

MaintenancePlan: This class represents the concept of a road maintenance plan. Most public road authorities will not allow utility owners to cut their newly paved road segment, which forms a constraint the manager scheduling the utility activities needs to consider. A road maintenance plan object can have many resurface objects, which represent the resurface activities and will be described next. The "MoratoriumMonths" attribute is used to store how long the road segment cannot be cut. The value of this attribute should be applicable to all resurface activities of the plan.

Resurface: This class represents the concept of a resurface activity. In our proposed model, the resurface activity pertains to a certain road segment; hence the influential area of a resurface activity must reside within the boundary of the corresponding road segment. The "Polygon" attribute denotes the influential area of a resurface activity. The "Duration" attribute records the start and end dates of a resurface activity. Note that a resurface activity can be a past, present, or future event.

UtilPart: This class represents the abstract concept of a utility facility. This class is a base class, and three classes, i.e., UtilFacility, UtilLine, and Abandonment, inherit from it. Basically, a road segment may have zero to many utility facilities, and the "Polygon" attribute depicts the boundary of a utility facility. The "BeginDate" attribute records the date when this facility becomes operational.

UtilFacility: This class represents the concept of an actual utility facility. Because it is derived from the "UtilPart" class, only the "Function" attribute and the "Clearance" attribute are defined. A utility facility object can be used to depict any facility or appurtenance with a clearance distance and must be contained in a utility section object.

UtilLine: This class represents the concept of an actual utility pipeline. Because it is derived from the "UtilPart" class, only the "Clearance" attribute is defined. A utility line object can be used to depict any pipeline with a clearance distance and must be contained in a utility section object. Note that a utility facility object, like a pole, can have two or more utility line objects, like a power line and a communication line suspended in the same pole.

Abandonment: This class represents the concept of an actual abandoned utility facility. Because it is derived from the "UtilPart" class, only the "Name" and "Reason" attributes are defined. An abandonment object can be used to depict any utility facility or pipeline that still exists in the field but no longer provides any service. An abandonment object must be contained in a utility network object.

UtilSection: This class represents the concept of a utility section, which contains a set of utility part objects.

The "Line" attribute is a "linestring" type in PostGIS, which describes a utility section as a set of lines. A utility section object must be contained in a utility network object and may contain zero or more utility permits.

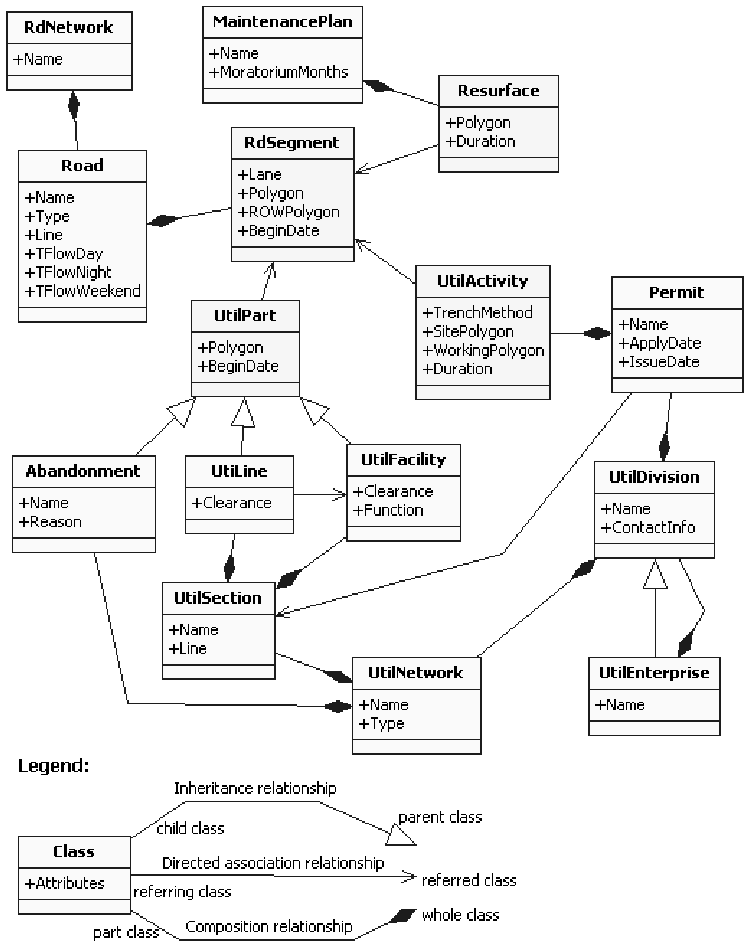


Figure 2. Model for spatiotemporal road and utility information

UtilNetwork: This class represents the concept of a utility network. A utility network object is designed to group utility section objects. For instance, in a city, the water utility network can be divided into many utility sections. Each section can be represented by a utility section object. The "Type" attribute is used to record the type of the utility network, e.g., water, wastewater, power distribution, gas, telecommunication, etc. The "Name" attribute records the name of the network.

UtilDivision: This class represents the owner of a utility network. It includes a person's contact information so that public road authorities can get in touch with him or her when any emergent event on the utility network occurs. A utility division object manages one or more utility networks. The "Name" attribute records the name of the organization.

UtilEnterprise: This class represents the concept of the utility division's business entity and inherits from the "UtilDivision" class. For example, a city government may be responsible for the water supply system, the wastewater treatment system, and the power distribution system. There must be three engineering divisions in the city government, so "UtilEnterprise" represents the city government whereas "UtilDivision" represents each engineering division. The Enterprise-Division structure can be easily extended to model any infrastructure owners' organizations.

Permit: This class represents the concept of a utility permit. A utility division may apply for several utility permits for each of their utility work. Similarly, a utility permit may pertain to one or more utility divisions who are all responsible for the utility work. A utility permit object can contain many utility activity objects. The two date attributes record when to submit the permit application and when to receive the approval.

UtilActivity: This class represents the concept of a utility activity. Each utility activity resides in one road segment. The "Duration" attribute records the duration of a given utility activity. The "TrenchMethod" attribute can be "trench," "trenchless" or "pole." The "SitePolygon" attribute contains the geometry of the trench or activity, whereas the "WorkingPolygon" attribute contains the geometry of the area needed for the work.

It should be noted that the "TranDate" attribute is added to each class in the proposed model because it represents the time when a given object is updated. The transaction time dimension is applicable to each class, whereas the valid time dimension is applicable to some of the classes in the proposed model with two formats. The first format of the valid time dimension is represented as the "BeginDate" attribute, which exists in the "RdSegment" and "UtilPart" classes. Since the road segments and the utility lines and facilities will not be demolished, the "EndDate" attribute of the valid time dimension is meaningless for the two classes. However, the "Duration" attribute of the "Resurface" and "UtilActivity" classes actually uses the start and end date fields to record the time period of a given activity.

6. Exploiting the Model

In this section, several hypothetical examples to exploit the capabilities of the model are discussed. The SQL language with system-defined or our customized spatiotemporal functions is used to explain how a spatiotemporal query can be done.

Finding Any Spatiotemporal Activity Violating the Constraints: As noted before, generally two or more utility activities performed by different utility owners cannot happen in the same place during the same period of time, unless they have been well coordinated in advance. Assume our database has accumulated enough road and utility infrastructure baseline data and recorded many utility permits data. Among these permits, it is difficult and time-consuming for a project manager to manually check all utility activities that violate the above constraint. Additionally, the pavement moratorium constraint of a road segment, the clearance distance of a utility facility or pipeline, and the traffic condition of a road all need careful examination before a utility permit can be issued. In SOD, the problem can be solved by the following steps:

1. Let Set1 = select * from UtilActivity where Duration.EndDate >= NOW
// find the utility activity data set that the work items have not been finished

2. Let Set2 = select * from Set1 as a, Set1 as b where st_intersect(a.Polygon, b.Polygon) is not null and a <> b
// list any pair of two objects that constitute a shared polygon, which means these two utility activities share the same place. Note that st_intersect is an OGC-defined function

3. Let Set3 = select * from Set2 where t_overlap(a.Duration, b.Duration) and a <> b
// because Set2 means the two activities occupy the same place, Set3 finds out whether their time durations are overlapped. Note that t_overlap is a customized function

4. Let Set4 = select * from Set1 join RdSegment join Resurface where st_intersect(Set1.Polygon, Resurface.Polygon) is not null and t_overlap(Set1.Duration, (NOW, Resurface.Duration.EndDate + Resurface.MaintenancePlan.MoratoriumMonths))
// if the space needed by Resurface and Set1 is the same, and if the work duration is overlapped, Set4 stores these records

5. Let Set5 = select * from Set1 join RdSegment join UtilPart where st_intersect(Set1.Polygon,

expand(UtilPart.Polygon, Clearance)) is not null
// if the space needed by expanding the clearance distance of a utility facility and Set1 is the same, Set5 stores these records

6. Let Set6 = select * from Set1 join RdSegment join Road where TFlow > certain_traffic_flow_valuee
// if the space needed by Set1 is the road with heavy traffic flows, Set6 stores these records

The result set is (Set3 union Set4 union Set5 union Set6), which means these planned utility activities violate one of the constraints.

Listing the Utility Activities That Can Be Performed Consecutively: In order to minimize the unnecessary pavement utility cuts, the project manager would like to list the utility activities that are planned to be performed in the approximate same place but at different times. These activities can be rescheduled to be performed consecutively to reduce unnecessary pavement utility cuts. Assume the minimum distance between the two activities that can be combined to be performed jointly is d. Assume "consecutively" means the difference between the end date of one activity and the start date of another consecutive activity is n days or less. In SOD, the problem can be solved by the following steps:
1. Let Set1 = select * from UtilActivity where Duration.EndDate >= NOW
// find the trench data set that the work items have not been finished

2. Let Set2 = select * from Set1 as a, Set1 as b where st_dwithin(a.Polygon, b.Polygon, d) and a <> b
// list any pair of two objects located nearby, which means these two utility activities will be performed in the approximate same place. Note that st_dwithin is a OGC-defined function

3. Let Set3 = select * from Set2 where (t_overlap(a.Duration, b.Duration) or a.Duration.EndDate + n >= b.Duration.StartDate) and a <> b
// Set3 finds out whether their durations are overlapped or very close

4. Let Set4 = Set2 – Set3
// the remaining data items are those activities that can be rescheduled. Set 4 is the result set

A simulation tool is currently under development. The tool will use the proposed model to simulate events of each utility activity in a small city in Taiwan. The water supply system, power distribution lines, communication lines, and natural gas pipelines are depicted in the tool. Their baseline and interdependency data are recorded in the database. With interdependency data in the database, users of the tool can find out the useful information to rescheduling.

7. Conclusions

With the ever-increasing demand for a streamlined analysis of utility activities to reduce unnecessary pavement utility cuts, public road authorities are making substantial efforts to improve the decision-making process regarding how to optimize the work plan. Because an information model is fundamental to be used to calculate and reschedule the utility activities, and because modern database technologies have provided powerful spatial query capabilities, the time dimension for the utility activities scheduling such as the one in this research can help project managers retrieve relevant information on demand. This research has designed a UML-based model that can describe spatiotemporal information with constraints. Further implementation and evaluation of the proposed model proposed is needed in order to demonstrate how such information technology can help reduce unnecessary pavement utility cuts.

References

[1] Chou, C.C., Caldas, C.H., and O'Connor, J.T. (2007). "Decision support system for combined transportation and utility construction strategy." Transportation Research Record: Journal of the Transportation Research Board, 1994, 9-16.

[2] Guting R.H., and Schneider, M. (2005). Moving Objects Databases, Morgan Kaufmann, San Francisco, CA.
[3] Jensen, K.A., Schaefer, V.R., Suleiman, M.T., and White D.J. (2005). "Characterization of utility cut pavement settlement and repair techniques." Proceedings of the 2005 Mid-Continent Transportation Research Symposium, Iowa State University, Ames, Iowa.
[4] Khogali, W.E.I., and Mohamed, E.H. (1999). "Managing utility cuts: issues and considerations." APWA International Public Works Congress, NRCC/CPWA.
[5] Lin, J.D. (2005). The M-Taiwan Project, Department of Transportation, Taiwan.
[6] Wilde, W.J., Grant, C., and White G.T. (2003). "Controlling and reducing the frequency of pavement utility cuts," Transportation Research Board Annual Meeting, Washington, D.C.
[7] Tansel, A., Clifford, J., and Gadia, S. (1993). Temporal Databases: Theory, Design, and mplementation, Benjamin-Cummings, San Francisco, CA.
[8] Wolfson, O., Xu, B., Chamberlain, S., and Jiang, L. (1998). "Moving objects databases: issues and solutions." Proceedings of the Tenth International Conference on Scientific and Statistical Database Management, IEEE, Capri, Italy, 111-122.
[9] Wolfson, O. (2002). "Moving objects information management: the database challenge." Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 2382, 15-26.

Social Network Analysis of Collaborative Entries for Construction Firms in International Construction Projects

Heedae Park[1], Woo Yong Jeong[2] and Seung Heon Han[3]

[1]Department of Civil Engineering, Yonsei University, Seoul, Korea. PH (822) 2123-2799; FAX (822) 364-5300; e-mail: parkheedae@yonsei.ac.kr
[2]Department of Civil Engineering, Yonsei University, Seoul, Korea. PH (822) 2123-2799; FAX (822) 364-5300; e-mail: trust-jung@gmail.com
[3]Associate Professor, Department of Civil Engineering, Yonsei University, Seoul, Korea. PH (822) 2123-2799; FAX (822) 364-5300; e-mail: shh6018@yonsei.ac.kr

Abstract

The current global construction market is growing tremendously largely due to globalization of world construction markets, rapid development of world-wide communication technologies, extensions of regional Free Trade Blocks, just to name a few. Korean companies, in search for a growth from the recession of domestic construction market, have expanded into the globalized market over the last decades. During this process, collaborative entries have increased remarkably to mitigate the burdens of entry risks. This study focuses on the formation of collaborative networks of Korean contractors when entering foreign markets. Social network analysis (SNA) is introduced to investigate a variety of collaboration patterns and also their impact on performance. To this end, real collaboration cases of 389 overseas projects executed by Korean companies since 1990 were collected and classified into firm's size, collaboration modes, and performance levels. Results of SNA showed a series of apparent tendencies of developing collaboration networks, in order to realize a better performance under the risky project conditions. This study is expected to be used to search a capable partner to gain improved outcome while considering the relevant network hubs as well as their risk levels.

Introduction

Although the current state of world economy causes increasing concerns due to the recent depressed real estate market, financial crisis, and a declined investment on infrastructure, volume of the global construction market has shown a truly remarkable growth over the last decades. In the wake of a global construction market, an increasing number of construction companies have expanded into the global market; not only large companies (LCs) but also small and medium-sized companies (SMCs). Likely, for Korean companies, long-lasting stagnation of domestic construction market leads them to expand into the global market as a way of balanced growth. However, international construction projects involve entry barriers and loads of uncertainties such as political, economic, cultural, legal, and technical risks. These risks can pose a significant threat such that even an internationally-experienced company fails to earn a decent level of profit from risky attempts of overseas projects. In this situation, collaborative entries are remarkably increasing to mitigate the burdens of entry risks that cannot be properly avoided when entering as a single company. Therefore, strategic collaborative formations and joint ventures among companies are on the increase although motives of collaboration modes are different. For example of Korean companies, international contract awarded by collaboration scheme established within LCs is increased from 4 projects in 1996 to 19 projects on 2006. Likewise, those between LC and SMC, and even within SMCs are also continuously increasing from 5 and 0 projects in 1996 up to 73 and 13 projects in 2006, respectively (ICAK, 2007).

With an increase of collaborative entries, previous research on collaboration in construction domain has been widely conducted, mainly focusing on the theoretical principles of partnering and alliancing, and identification of critical factors (Alarcon and Mourgues, 2002; Anvuur and Komaraswamy, 2007; Wong and Cheung, 2005). However, there is little research on collaboration patterns from the network perspective. This study focuses on the formation of collaborative networks between Korean contractors when entering new foreign markets. Social network analysis (SNA) is introduced to investigate a variety of collaboration patterns and also their impact on performance. Accordingly, collaboration cases of 389 international construction projects executed by Korean companies from 1990 to 2006 are collected. These cases are then classified into firm's size, collaboration modes, and performance levels using SNA approach. Finally, this study intends to provide a useful direction in establishing business network strategies when forming cooperation with either internationally-experienced or inexperienced companies through collaboration network analysis.

Research Background

There are a number of researches on firms' collaboration in the construction industry. Anvuur and Kumaraswamy (2007) presented critical influence factors for successful partnering and alliance between construction firms. Wong and Cheung (2005) quantitatively analyzed relationship of factors for partnering through structural equation modeling. In addition, Alarcon and Mourgues (2002) presented success/failure factors of contractor selection in partnering. Lin (2003) also investigated how effectively general contractors can improve relationship with subcontractors. Although these previous researches provided useful knowledge and principles of diverse forms of collaboration, they had limitations in identifying the common aspects and similar patterns from macro-industrial level and network perspective. Therefore, it makes exploring the collaboration networks between firms worth studying to investigate their synergy effects and establish better strategies for fostering collaboration.

Social network analysis (SNA) is a methodology to identify social structures through analyzing interactions and interrelationship of a set of actors (Hu and Rachera, 2008). SNA focuses more on relational properties among components rather than individual attributes. Accordingly, the main objective of SNA is in identifying and analyzing conditions of social relationships' among components (Nooy et al. 2005).

SNA has been applied to investigate various relationships among actors and organizations, and knowledge diffusion in the field of social sciences and economics. Lee (2005), for example, compared development patterns between the automobile industry and semiconductor industry from the network perspective. Hu and Rachera (2008) presented knowledge diffusion network by analyzing co-authorship in the hospitality research domain. Recently, SNA research in the construction field has gradually increased, as construction projects have a wide range of participants that are mutually interrelated and cooperated. Particularly, Pryke (2004) investigated managerial attributes of UK construction projects in connection with procurement modes, supporting the power of SNA application in construction domain. Chinowsky et al. (2008) also proposed an SNA approach to project management that can be used to integrate classic project management concept and knowledge sharing perspective for achieving a higher project performance.

Data Collection and Assumptions

This study collected international project cases collaboratively performed by Korean companies from 1990 to 2006. Those projects executed by a single entity or contracted for design service were excluded from the analysis; thus, 389 collaborative projects were further analyzed. To categorize the collaboration forms, the authors first used the guideline for distinguishing LCs and SMCs, enacted by the Korean government. According to the criterion, LCs should employ more than 300 full-time workers and hold more than $ 2 million capital. Following table represents the summary of collected data in line with collaboration modes: (1) within LCs, (2) between LC and SMC, and (3) within SMCs. Besides, sample projects are also classified according to the conditions of contracts: joint venture (JV) and prime contractor vs. subcontractor.

Table 1. Summary of Analysis Sample

	Joint Venture (JV)	Prime contractor vs. Subcontractor
LCs	10	76
LC and SMC	19	231
SMCs	3	50
Total	32	357

This study adopts Pajek® SNA program that enables to visualize and analyze large networks having thousands of vertices. Since Pajek® classifies network and color vertices by aid of auxiliary network, the authors discerned profit performances of all projects into dichotomy zones: gain or lose, and similarly, labeled blue and yellow color to LCs and SMCs, respectively. Since this study focuses on the collaboration within Korean companies rather than that with foreign or host country' local companies, this study limited collaboration motives, primarily focusing on the familiarity and mutual-trust between home country's companies, rather than concentrating on the mitigation of entry barrier, reinforcement of bid qualification, and utilization of edged technologies obtainable from foreign firms.

Descriptive Analysis

Figure 1 shows the collaboration network of 389 entire projects completed by 136 companies. Each circle indicates an actor (contractor), while the line between actors shows an existence of collaboration between two actors. The bigger a circle is, the more a company participates in international projects. The direction of an arrow shows contractual relationships between two actors. If an arrow points B from A, then B implies a subcontractor of A. If there is a parallel contract such as a joint venture, the arrows point to both sides. As shown, actors that have suffered loss are positioned at left side, and vice versa. The brightness of circles classifies LCs (dark actors) and SMCs (bright actors).

In the entire collaboration network, the majority of lines are joined into several actors such as H11, S4, and D7. Interestingly, whereas H11, J8, and D7 in LCs group and their collaborative contractors have realized profits together, D23, S4, and their collaborators have suffered losses. It implies that we can distinguish promising networks from negative ones to achieve a better performance. Thus, it is possible for a particular company to search for probable partners, because good-quality networks, in general, continue to result in respectable performance, and vice versa.

In more details, Figure 2 shows the collaboration network within LCs. Among the actors, five LCs formed the loss-making set but they participated in few projects. There are two types of contractors in the LCs network. Firms such as H11, H15, and S4 collaborated with many companies, whereas other firms including H3, G3 and A20 cooperated with only one company. Interestingly, the collaboration network within LCs pinpoints that more internationally-experienced companies gained higher profits and subsequently they continued to closely collaborate within themselves. For example, H11 and D7, who are the most active companies in this network, gained 5.4%, and 5.5% of average profit, respectively, posing a higher average profit compared to 1.9% of other companies. It was found that the two companies continued to collaborate with internationally-experienced firms or allied enterprises belonged to their large conglomerates.

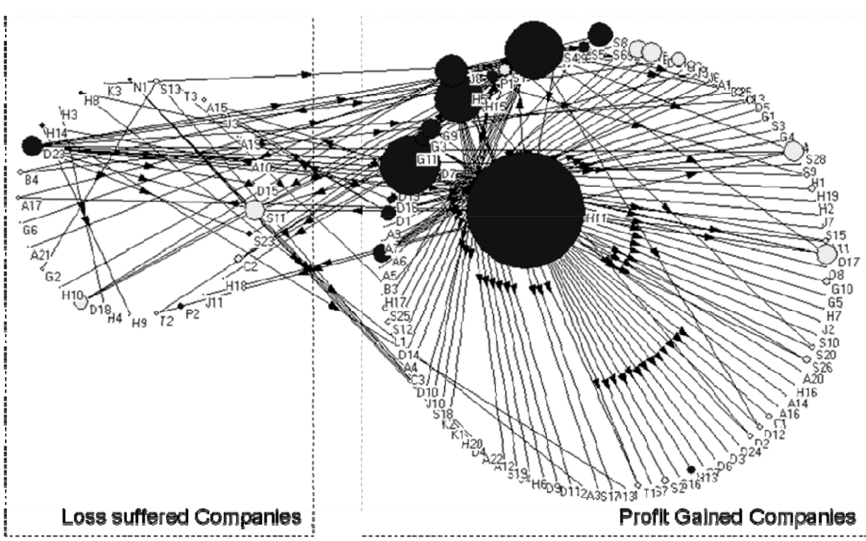


Figure 1. Collaboration Network (All)

The collaboration network between LCs and SMCs shows the most complex relationships where numerous companies work together and collaborate more frequently. In this scheme, the majority is the collaboration established between prime contractors (LC) and subcontractors (SMC), occupying 231 out of 250 cases. Only 19 projects were performed through a form of JV where they achieved relatively higher profit than others. It was found that SMCs in this collaboration were chosen as JV partners because they possessed competent skills in construction methods and equipments for building special facilities including power plant, electronic, or subway foundation.

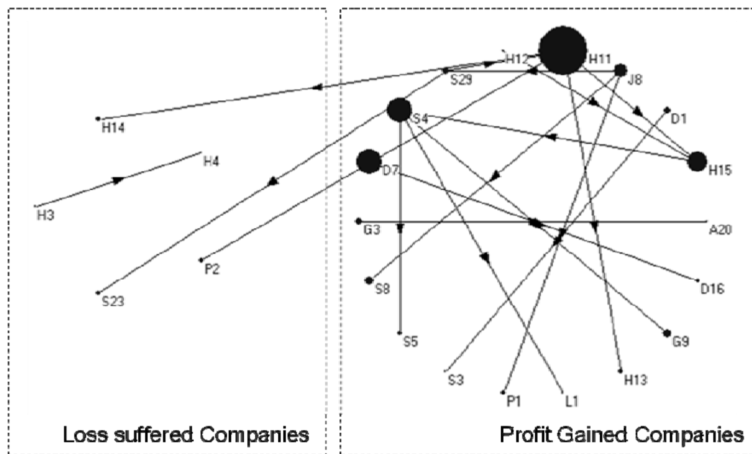


Figure 2. Collaboration Network (between LCs)

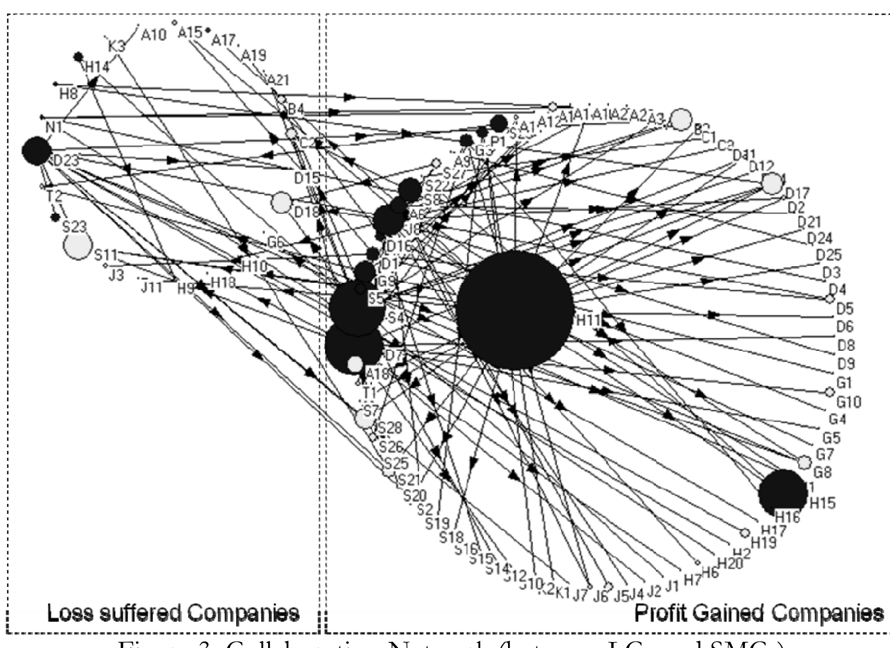


Figure 3. Collaboration Network (between LCs and SMCs)

Meanwhile, SMCs that have collaborated continuously with the same LC gained higher profits than other cases. On the other hand, LCs gained more profit when they have collaborated with many pools of SMCs. More specifically, in the case where LCs collaborated with more than five SMCs, they gained 3.69% profit on the average, while other cases obtained only 2.62% profit. This partly explains why LCs prefer to choose SMCs proposing a lower cost among potential bidders rather than engaging a particular subcontractor in their networks. It is fairly paradoxical to the fact that SMCs could gain higher profits through collaborating with a small number of LCs under a quasi-hierarchy system or with a long-term relationship with a particular LC.

Finally, the collaboration network established within SMCs is more segmented than other networks (see Figure 4). Except the few cases of S13, S1, and H5, the most actors were related with only one contractor, implying that they could not maintain capturing project opportunities with this weak network. In another aspect, the larger SMCs such as D18, S28, and S21 have obtained many contracts but worked in cooperation with only one company. This is likely because those companies achieved competitive positions with collaborating each other, but it also shows the limitation of SMCs network in developing more capable relationships when entering the international market.

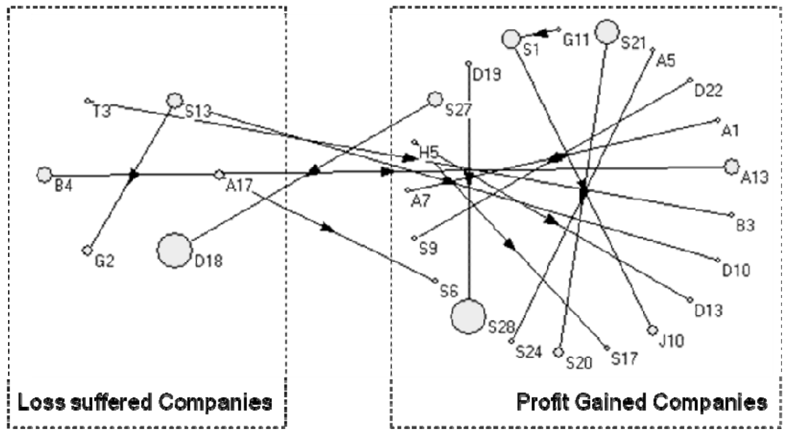


Figure 4. Collaboration Network (between SMCs)

Network Attributes

SNA displays not only a visual representation but also network variables that express structural attributes, which in turn helps analyzing a network in further details. This paper investigates degree, density, and centrality through examining network variables. Also, several other properties are investigated for comprehensive understanding of networks.

Degree, Density, and Centrality

Degree indicates the number of actors collaborating with a specific actor, and it is a good index to measure regional centrality. Density represents a level of crowd; thus, is calculated by dividing the number of existing lines with the number of all possible lines, therefore, density varies from 0 (no connection in the network) to 1 (every vertex are interrelated). Lastly, centrality denotes the grade of structural efficiency. There are three types of centrality whether the focus is placed on inward directive, outward directive, and total connection : indegree centrality, outdegree centrality, and total centrality. Equations of three attributes above are like following.

Degree = number of connected vertices to a vertex

Density = l/(n(n-1)) (l = number of existent lines, n = number of existent vertices)*

Centrality = (variation in the degree of vertices)/(maximum degree variation)

Table 2 shows the summary of each networks' attributes. Despite the low level of density as a whole, the networks established within both LCs and SMCs are more intense, because density is inversely proportional to a network size. Meanwhile, networks within LCs, and between LC and SMC have higher centrality, particularly showing a tendency of outdegree centrality. This implies that there exist a small number of large contractors connected with many subcontractors such that networks tend to migrate toward several hubs. By contrast, the centrality of SMCs network shows the lowest value in all, as the SMCs' network is fairly segmented and scattered.

Frequency Distribution and Structural Equivalence

This paper analyzes two more attributes: frequency distribution and structural equivalence. The former refers to a summary of social oftenness to extend a cooperation cluster, and the latter represents the hierarchical formation of relationship among all actors in a network. As shown in Table 3, most contractors, around 52%~78%, are found to be collaborated with only one contractor in every network. The most sociable contractors within LCs, and between LC and SMC built up huge clusters engaged with 18 and 28 contractors, respectively, whereas that of SMCs network constitutes a small cluster with only 3 contractors. However, as previously mentioned, there is a weak evidence to connect between how sociable a contractor is and whether a contractor could make profits. It rather relies on how experienced a companion contractor is and whether a contractor establishes a collaborative network with a firm that has demonstrated a better performance.

Table 2. Summary of Network Attributes

	Degree (min - max)	Density	Centrality			% of firms gaining profit		
			In	Out	All	LCs	SMC	All
LCs	1 - 18	0.06	0.05	0.15	0.08	78.3	-	78.3
LC and SMC	1 - 28	0.01	0.02	0.20	0.09	69.6	80.2	77.9
SMCs	1 - 3	0.02	0.01	0.05	0.01	-	77.8	77.8

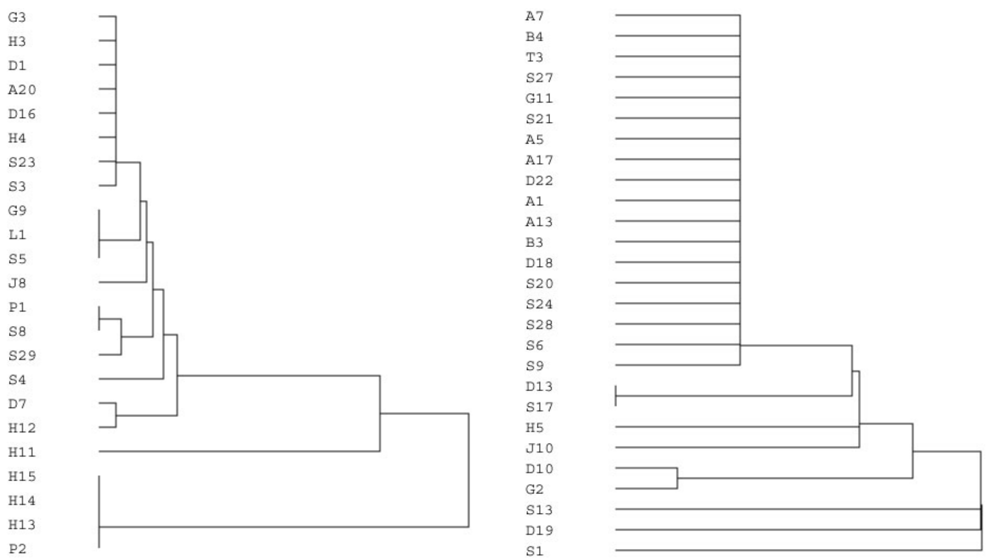
Table 3. Frequency Distribution of Cluster Numbers

Cluster	LCs		LC and SMC		SMCs	
	Freq.	Cum.Freq.%	Freq.	Cum.Freq.%	Freq.	Cum.Freq.%
1	15	65.2	55	52.9	21	77.8
2	-	65.2	19	71.2	3	88.9
3	5	86.9	11	81.7	3	100.0
4	1	91.3	4	85.6	-	
5	-	91.3	2	87.5	-	
⋮	⋮	⋮	⋮	⋮	-	
14	1	95.7	-	94.2	-	
⋮	⋮	⋮	⋮	⋮	-	
18	1	100.0	-	96.2	-	
⋮	-		⋮	⋮	-	
21	-		2	99.0	-	
28	-		1	100.0	-	
Sum	23	-	104	-	27	-

Figure 5(a) and Figure 5(b) represent the structural equivalence of LCs network and SMCs network, respectively. Diagram of network between LCs and SMCs is not included here for the sake of brevity. Since the SMCs' network is broken down into lots of pairs, it has a simple and uniform structure, while those within LCs, and between LC and SMCs have perpendicular and hierarchical structure. Based on the analysis results, the authors conclude that searching an experienced companion in the LCs' network is relatively easy, but it is not the case in the SMCs' network because capable contractors are not clearly disclosed in this type of system's structure.

Conclusion

This study focuses on the formation of collaborative networks established within Korean companies when expanding into new foreign markets. The authors analyzed 398 cases of international projects executed by such collaborative organizations as JV and subcontracting. SNA was introduced to investigate a variety of collaboration patterns and also their impact on the level of performance. In this process, the authors attempted to relate the networks with their profit performances, in association with collaboration modes. Results of SNA emphasized that; (1) within the LCs' network, making continuous collaboration with a particular company intended to result in higher profits, (2) between LCs and SMCs network, SMC could realize a higher profit when making a continuous collaboration with the same LC, whereas LCs gained a less profit when working with one particular SMC, (3) if SMCs failed to achieve a competitive edge in international construction per sea, they were difficult to find a promising network and so to gain a better profit, and (4) the more hierarchical the structural equivalence, the higher the possibility for searching excellent partners to gain improved profit. Recalling that the main scope of this research is to analyze collaborations made within home country's companies, future procedural research will focus on more broadened areas including diversities of collaboration patterns, regional distributions, and project types, in order to explore and understand how these factors influence network conditions and performance. Through an extension of current analysis, construction firms can effectively choose a suitable entry partner to capture more overseas project opportunities and better results while considering the relevant network hubs as well as their risk levels.



(a) (b)

Figure 5. Structural Equivalence ((a): LCs, (b) SMCs)

Acknowledgement

This research was supported by a grant (07 Urban Renaissance B03) from High-Tech Urban Development Program funded by the Ministry of Construction and Transportation of Korea.

References

[1] Alarcon, L. F. and Mourgues, C. (2002). "Performance Modeling for Contractor Selection." J. Manage. Eng., 18(2), 52-60.
[2] Anvuur, A. M. and Kumaraswamy, M. M. (2007). "Conceptual Model of Partnering and Alliancing." J. Const. Eng. and Manage., 133(3), 225-234.
[3] Chinowsky, P., Diekmann, J., and Galotti, V. (2008). "Social Network Model of Construction." J. Const. Eng.and Manage., 134(10), 804-812.
[4] Hu, C. and Rachera, P. (2008). "Visual Representation of Knowledge Networks: A Social Network Analysis of Hospitality Research Domain." Int'l J. Hospitality Manage., 27, 302-312.
[5] International Contractors Association of Korea (ICAK) (2008) "International Construction Information Service: Summary of the present state."<http://www.icak.or.kr/sta/sta_0101.php> (Dec. 8, 2008).
[6] Lee, G. H. (2005). "Mutual Influence of Corporate Networks and Industrial Change." J. of Korea Society of Social Science, 39(4), 1-40.
[7] Lin, P. H. (2003). "Decision Support for Subcontracting Procurement Based on Multi-Attribute Utility Theories." Const. Research Congress 2003, Honolulu, Hawaii, ASCE.
[8] Ministry of Land, Transportation, and Maritime Affairs (2007). "Development of Pilot Program for Collaboration of Large Enterprise and Small-Medium Enterprise." Research Report, No. 2007-03.
[9] Nooy, W., Mrvar, A., and Batageli, V. (2005). "Exploratory Social Network Analysis with Pajek." Cambridge University Press, Cambridge, UK.
[10] Pryke, S. D. (2004). "Analysing construction project coalitions: exploring the application of social network analysis." Constr. Manage. and Economics, 22(8), 787-797.

Prediction of Diaphragm Wall Deflection in Deep Excavations Using Evolutionary Support Vector Machine Inference Model (ESIM)

Min-Yuan Cheng[1] and Yu-Wei Wu[2]

[1]Professor, Dept. of Construction Engineering, National Taiwan University of Science and Technology, Taiwan; POST: #43, Sec. 4, Keelung Rd., Taipei, Taiwan, R.O.C. 106, Department of Construction Engineering; e-mail:myc@mail.ntust.edu.tw.
[2]Ph.D. Candidates, Dept. of Construction Engineering; National Taiwan University of Science and Technology, Taiwan; PH (886) 227330004; e-mail:D9305503@mail.ntust.edu.tw.

Abstract

Problems in deep excavations are full of uncertain, vague, and incomplete information. In most instances, successfully solving such problems depends on experts' knowledge and experience. The primary object of this research was to propose an "Evolutionary Support Vector Machine Inference Model (ESIM)" to predict wall deformation in deep excavation in Taipei Basin. ESIM is developed based on a hybrid approach that fuses support vector machines (SVM) and fast messy genetic algorithm (fmGA). SVM is primarily concerned with learning and curve fitting; and fmGA with optimization. Fifty-seven wall deformation monitoring database were collected based on monitoring data and compiled from prior projects. Fifty-two of 57 were selected for training, leaving 5 valid cases available for testing. Results show that ESIM can successfully predict the deflection and apply to contractors utilizes the knowledge and experience from past projects to predict wall deformation of new projects. Therefore the construction and foundation construction contractors can update wall deflection monitoring data of different stages during deep excavation process, in order to predict the wall deformation of the next stage and examine whether the max deflection is within the controlled range. The results are used as guidelines on site safety and risk management.
Keywords: SVM, fmGA, Diaphragm Wall; Excavation; Deflection

1. Introduction

Braced diaphragm wall structures are commonly used in deep excavation projects to improve the safety and quality of construction. Therefore, how to use monitored data effectively to predict diaphragm wall deflection, ensure project safety and prevent costly damage represents a critical issue. Data on diaphragm wall deflection is regularly monitored to ensure construction quality and the safety of adjacent buildings - particularly in high density urban settings. However, the complexity of geotechnical parameters and variety of construction factors make the behavior of the soil/wall/prop structures difficult to determine. Peck (1969), Goldberg et al. (1976), Long (2001) have previously identified the key factors in deep excavation to include soil type and properties, excavation depth, and wall stiffness, among others. The first task for this study was to compile historical data from relevant and reliable deep excavation cases. Afterward, approaches to estimate retaining wall support system defection, e.g., finite element analysis, were evaluated and applied.

Finite element analysis has previously been employed to simulate the braced diaphragm wall system (Clough and Hansen 1981; Powrie and Li 1991). However, results are heavily dependent upon the constitutive behavior of soil. As model parameters are usually obtained from laboratory tests, they are unable to fully represent in-situ soil properties due to sample disturbance, in-situ environmental conditions, the diverse effects of construction, and so on. Feedback analysis is commonly applied to field measurements to determine soil parameters (Gioda and Sakurai 1987). Whitted et al. (1993) applied finite element analysis to model the top-down construction of a seven-story, underground parking garage at Post Office Square in Boston. By using optimization approaches, factors were modified to improve agreement with the measured data without recourse to parametric iteration. Ou and Tang (1994) proposed a nonlinear optimization technique to determine soil parameters for deep excavation finite element analysis and studied a hypothetical excavation case under a variety of ground conditions. Chi et al. (2001) obtained optimized parameters by applying an optimization technique for back-analysis that produced results in good agreement with field measurements.

However, these approaches presented several important difficulties which rendered them inadequate for general application. The construction industry is replete with myriad uncertainties that make management exceedingly complex. Various scientific and engineering specializations have been paying increasing attention in recent years to the fusing of different artificial intelligence (AI) paradigms to achieve greater efficacy in results. A number of studies have demonstrated that performances achieved by fusing different AI techniques are better than those achieved by employing a single conventional technique (Yang and Yau 2000). Fast messy genetic algorithms (fmGA) and the support vector machine (SVM) are two tools that have been applied successfully to solve various construction management problems. Considering the characteristics and merits of each, this paper combines the two to propose a new Evolutionary Support Vector Machine Inference Model (ESIM) (Cheng and Wu 2008). In the ESIM, the SVM is employed primarily to address learning and curve fitting, while fmGA addresses optimization. This model was developed to achieve the fittest C and gamma parameters with minimal prediction error. This study applied diaphragm wall deflection data previously compiled from 18 metropolitan Taipei projects to the ESIM to prediction result accuracy.

2. The Evolutionary Support Vector Machine Inference Model (ESIM)

Support vector machines and fast messy genetic algorithms represent recently developed AI paradigms. SVMs were first suggested by Vapnik (1995) and have recently been applied to a range of problems that include pattern recognition, bioinformatics, and text categorization. An SVM classifies data using different class labels by determining a set of support vectors that are members of the set of training inputs that outline a hyper plane in a feature space. It provides a generic mechanism that fits the hyper plane surface to the training data using a kernel function. The user may select a kernel function (e.g. linear, polynomial, or sigmoid) for the SVM during the training process, which identifies support vectors along the function surface. Using SVMs presents users with the problem of determining optimal kernel parameter settings. Therefore, obtaining SVM parameters must occur simultaneously. Proper parameter settings can improve SVM prediction accuracy, with parameters that should be optimized including penalty parameter C and kernel function parameters such as the gamma of the radial basis function (RBF) kernel. In designing an SVM, one must choose a kernel function, set kernel parameters and determine a soft margin constant C (penalty parameter). The Grid algorithm is an alternative to finding the best C and gamma when using the RBF kernel function. However, this method is time consuming and does not perform well (Hsu and Lin 2002; Huang, Wang 2006). Fast messy genetic algorithms were developed by Goldberg et al. in 1993. Unlike the well-known simple genetic algorithm (sGA), which uses fixed length strings to represent possible solutions, fmGA applies messy chromosomes to form strings of various lengths. Its ability to identify optimal solutions efficiently for large-scale permutation problems gives fmGA the potential to generate SVM parameters C and gamma simultaneously. Considering the characteristics and merits of each, this paper combined the two to propose an Evolutionary Support Vector Machine Inference Model (ESIM). In the ESIM, the SVM is employed primarily to address learning and curve fitting, while fmGA addresses optimization. This model was developed to achieve the fittest C and gamma parameters with minimal prediction error. ESIM structure is illustrated in Figure 1.

The following three steps must be followed to establish an accurate fmGA-based parameter optimization model:

(1) Establish an SVM training model. The SVM trains a prediction model using default parameters and a training dataset.

(2) Obtain average accuracy. A training dataset is used for each chromosome representing C and gamma to train the SVM and calculate accuracy. When said accuracy is obtained, each chromosome is evaluated using a fitness function.

(3) Set termination criteria. The process ends once termination criteria are satisfied. In the absence of such, the model will proceed to the next generation.

(4) Search fmGA parameters. The model searches for better solutions by genetic operations.

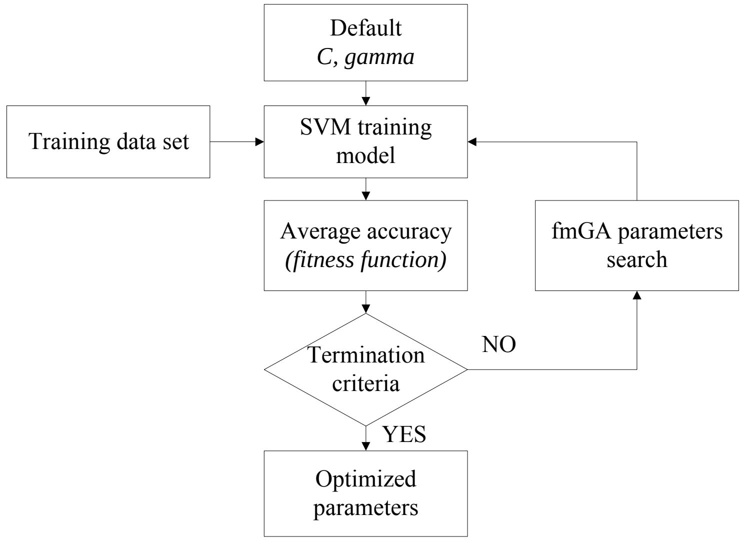


Figure 1. Structure of the ESIM.

3. ESIM for Predicting Diaphragm Wall Deflection

3.1 Knowledge Representation of Diaphragm Wall Deflection in Deep Excavation

Diaphragm wall systems are widely used in deep excavation, and significant amounts of data are collected to monitor their safety. As such large amounts of data have been accumulated, how to use such to improve the safety of current and future projects represents an important area of potential development. The ESIM has been adopted to solve this problem, employing historical data to predict diaphragm wall deflection during excavation. The key initial issue faced is how to configure data into a useable format. In Figure 2, W represents diaphragm wall thickness; H the temporary depth of excavation; Ri the observation point factor where 18 segments are set; and He the final depth of excavation. Embedment depth is typically set as 0.8 He. However, in cases where embedment depth is less than 0.8 He, deflection between the bottom of the diaphragm wall and 19th observation point is assigned as linear and converges to zero and the total depth of diaphragm wall is set as 1.8 He. Referring to Jan et al. (2002), seven important factors were selected as inputs and one output was set. Each observation point can be regarded as an individual case, with related parameters illustrated as follows:

<u>Seven Inputs:</u>

(1)Diaphragm wall thickness: W.

(2)Excavation depth: H.

(3)The equivalent SPT-N value between H+0.25He and H-0.25He: $\overline{N}$.

(4)The factor of an observation point factor linearly interpolated by the depth: R.

(5)The deflection of the observation point in the last stage, i.e., the (i-1)-th stage in the current i stage in excavation: $D_{i\text{-}1}$.

(6)The deflection of the observation point in the (i-2)-th stage: $D_{i\text{-}2}$.

(7)The deflection of the observation point in the (i-3)-th stage: $D_{i\text{-}3}$.

<u>One Output:</u>

(1)The deflection of the observation point in i-th stage: Di.

To prevent the absence of fifth to seventh inputs, i has to be greater than or equal to three. When i=3, the Di-3 is set as zero.

3.2 Historical cases

Eighteen historical cases from metropolitan Taipei, Taiwan were collected. These cases are listed in Table

1, which provide information on the number of excavation stages, excavation depth and construction method used (top-down or bottom-up). The number of stages in these cases varied from four to seven. As each stage was treated individually, these cases comprised 93 stages in total. Excluding the first and second stages of construction, 57 stages of valuable data were collected. The first seventeen construction cases, including 52 stages total, were used for training. The remaining five stages of the 18th case were employed in testing. Nineteen observation points were set, although excavation depths were not uniform. Therefore, 19 sets of data were collected in each stage. Based on the above, 52×19 = 988 training data sets and 5×19 = 95 testing data sets were collected.

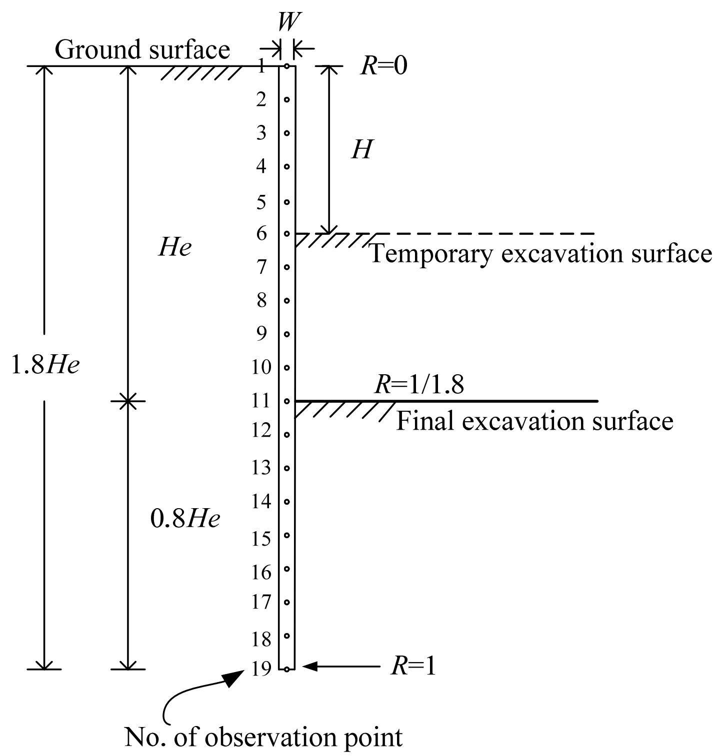


Figure 2. Representation of the Diaphragm Wall Structure

4. Comparison of Results

Training data (988 sets from 52 excavation stages) and testing data (95 sets from 5 excavation stages) were used to develop the ESIM diaphragm wall deflection prediction system. After training, the RMSE (Root Mean Square Error) equaled 3.794%. In Figure 3, the accuracy of maximum diaphragm wall displacements is demonstrated by comparing results with actual measurements and the average correlated coefficient (ACC) between the maximum predicted wall displacement and the maximum measured wall displacement (average of [predicted/measured]). ACCtraining equals 0.8983 and ACCtesting equals 0.8898. Among the 52 training excavation stages, there were 28 cases with relative errors less than 10%; 12 cases with relative errors between 10% and 20%; and 12 cases with relative errors exceeding 20%. If we define the criterion of failed prediction as an error of maximum predicted displacement that exceeds 20%, then 12 of 52 can be considered to have failed in prediction, i.e., the accuracy of diaphragm wall deflection prediction using this model is 76.9%. The data of project No. 18 (the project reserved for use in testing data) and its 5 stages with 5×19=95 sets of testing data were calculated and, while the same criterion was taken, 5 of the 5 were qualified. This gives an accuracy of prediction of 100%. To sum up training and testing data results, 12 of 57 sets of results fail to meet the criterion, i.e. the model achieves an accuracy of 78.94%. This result is an improvement one than done by Jan et al., which used NNs. In the following section, improvements will be applied to the prediction model to improve results even further.

Table 1. 18 Historical Excavation Projects in Metropolitan Taipei.

No.	Stages	Depth (m)	Construction method	No.	Stages	Depth (m)	Construction method
1	5	12.30	Top-down	10	6	14.05	Top-down
2	4	13.90	Bottom-up	11	4	13.60	Top-down
3	6	16.00	Top-down	12	5	17.35	Bottom-up
4	5	12.60	Top-down	13	5	13.15	Top-down
5	5	12.30	Top-down	14	5	23.85	Top-down
6	5	12.25	Top-down	15	6	19.40	Top-down
7	4	10.00	Top-down	16	6	19.40	Top-down
8	6	18.95	Top-down	17	5	13.70	Top-down
9	4	9.30	Top-down	18	7	19.70	Bottom-up

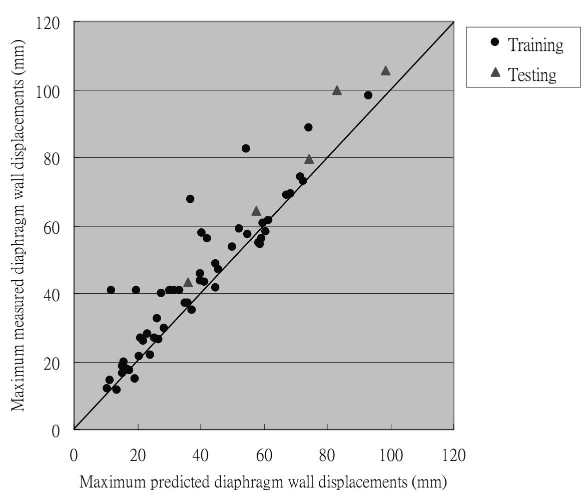


Figure 3. Measured vs. Predicted Maximum Diaphragm Wall Displacements.

4.1 An improvement for a new excavation project

The typical deep excavation project has many stages and the deflection observed in any given stage is highly correlated to deflection parameters in previous stages. Therefore, diaphragm wall deflection data from prior stages are important inputs to help predict the values of deflection variables in succeeding stages of an excavation project. As diaphragm deflections accumulate during an excavation, data from previous stages can be employed to predict deflection in the following stage with improved accuracy. Based on the above, project No. 18 data shown in Table 1 are treated as a new excavation project. In this project, the depth of the diaphragm wall is 35 meters and the total excavation depth is 19.7 meters. Seven excavation stages are adopted as follows: 1st stage: 2.8 meters; 2nd stage: 4.9 m; 3rd stage: 8.6 m; 4th stage: 11.8 m; 5th stage: 15.2 m; 6th stage: 17.3 m and 7th stage: 19.7 m. Monitored data from preceding stages can be adding into the training data as a new excavation project progresses from stage to stage.

For each excavation stage after the 2nd, data compiled from previous stages were added into the training data to present the individual characteristics of this particular project instantaneously. As shown in Table 2, errors have been reduced and accuracy improved by this modified process. The modified process significant improved ACCtesting compared to the previous result (from 0.8898 to 0.8927). Detailed results on wall deflection at every stage are shown in Figure 4. According to the results, the improvement works due to the adding of previous stages' data from the current project. Such data may be highly related with the prediction target based on a project's discrete characteristics.

Table 2. Results of the Modified Process Applied to the New Excavation Project

Excavation stage	Measured Max. displacement (mm)	Predicted Max. displacement (mm)	Original Error (%)	Modified Max. displacement (mm)	Modified Error(%)
3rd	43.44	35.84	**17.49**	35.84	**17.49**
4th	64.34	57.54	**10.57**	57.72	**10.29**
5th	79.57	73.88	**7.15**	73.89	**7.14**
6th	99.64	84.14	**15.56**	84.14	**15.56**
7th	105.72	101.14	**4.33**	102.42	**3.12**

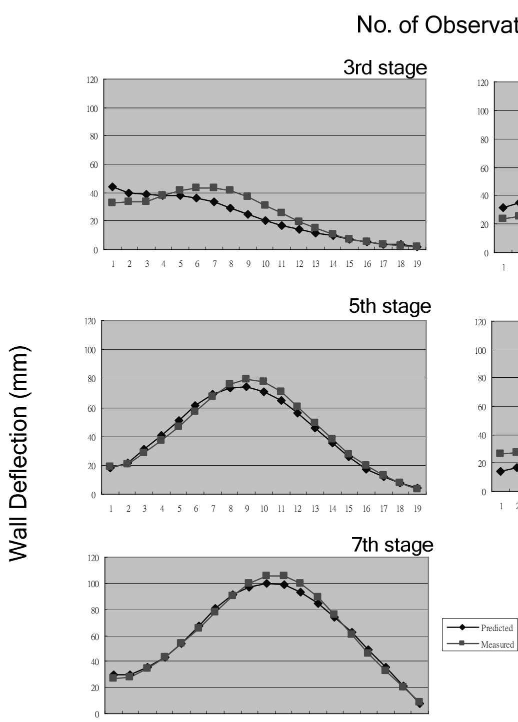


Figure 4. Wall Deflection Prediction Using the Modified Process

5. Conclusions

As useful information is hidden within monitored data, the ESIM may be employed to extract the critical effects of diaphragm wall deflection. Diaphragm wall deflection predictions not only employ historical case data, but also the data of previous stages in the training sets in order to reflect in-situ particularities. By applying ESIM, a strict understanding of parameters or their effects is not required. The magnitude of deflection and the position where the maximum displacement occurs in deep excavation diaphragm walls can, therefore, be predicted to ensure safety during the construction process. Deflection in the embedded position can also be performed. This permits engineers to make highly accurate appraisals of the diaphragm wall structure prior to starting the next excavation stage.

References

[1] A.J. Whittle, Y.M.A. Hashash, and R.V. Whitman (1993), "Analysis of deep excavation in Boston", Journal of Geotechnical Engineering, 119 (1), pp. 69-90.

[2] C.Y. Ou and Y.G. Tang (1994), "Soil parameter determination for deep excavation analysis by optimization", Journal of the Chinese Institute of Engineers, Transactions of the Chinese Institute of Engineers,Series A/Chung-kuo Kung Ch'eng Hsuch K'an, 17 (5), pp. 671-688. (In Chinese)

[3] C.-L. Huang, C.-J. Wang (2006), "A GA-based feature selection and parameters optimization for support vector machines", Expert Systems with Applications, 31, pp. 231-240.

[4] D.T. Goldberg, W.E. Jaworski, and M.D. Gordon (1976), "Lateral support systems and underpinning", Volume 1, Design and construction, Federal Highway Administration Report FHWA RD-75-128.

[5] G. Clough and L.A. Hansen (1981), "Clay anisotropy and braced wall behavior", Journal of the Geotechnical Engineering Division, 107 (7), pp. 893-913.

[6] G. Gioda and S. Sakurai (1987), "Back analysis procedures for the interpretation of field measurements in geomechanics", International Journal for Numerical and Analytical Methods in Geomechanics 11 (6), pp. 555-583.

[7] Hsu, C. W., Lin, C. J. (2002), "A simple decomposition method for support vector machine", Machine Learning, 46 (1–3), pp. 219–314.

[8] J.C. Jan, S.L. Hung, S.Y. Chi, and J.C. Chern (2002), "Neural network forecast model in deep excavation", Journal of Computing in Civil Engineering, 16 (1), pp. 59-65.

[9] M. Long (2001), "Database for retaining wall and ground movements due to deep excavations", Journal of Geotechnical and Geoenvironmental Engineering, 127 (3), pp. 203-224.

[10] Min-Yuan Cheng, Yu-Wei Wu. (2008), "Dynamic Prediction of Project Success Using the Evolutionary Support vector machine Inference Model", the 25th International Symposium on Automation and Robotics in Construction.

[11] R.B. Peck (1969), "Deep excavations and tunneling in soft ground", Proceeding 7th International Conference Soil Mechanics and Foundation Engineering, University Nacional Autonoma de Mexico Instituto de Ingenira, Mexico City, pp. 225–290.

[12] S.Y. Chi, J.C. Chern, and C.C. Lin (2001), "Optimized back-analysis for tunneling-induced ground movement using equivalent ground loss model", Tunnelling and Underground Space Technology, 16 (3), pp. 159-165.

[13] Vapnik, V.N (1995), "The nature of statistical learning theory", New York: Springer.

[14] W. Powrie and E.S.F. Li (1991), "Finite element analysis of an in situ wall propped at formation level", Geotechnique, 41 (4), pp. 499–514.

[1] Yang, J.B. and N.J. Yau (2000), "Integrating case-based reasoning and expert system techniques for solving experience-oriented problems", Journal of the Chinese Institute of Engineers, 23 (1), pp. 83–95.

Construction Scheduling Optimization by Simulated Annealing

M. König[1] and U. Beißert[2]

[1]Chair of Theoretical Methods for Project Management, Bauhaus-University Weimar, Marienstr. 7, 99423 Weimar, Germany, PH (49) 3643-584825; FAX (49) 3643-584565; e-mail: markus.koenig@uni-weimar.de
[2]Chair of Construction Management, Bauhaus-University Weimar, Marienstr. 7, 99423 Weimar, Germany, PH (49) 3643-584455; FAX (49) 3643-584565; e-mail: ulrike.beissert@bauing.uni-weimar.de

Abstract

The process of generating optimized schedules for construction projects is a very time-consuming. For each construction task several execution restrictions, conditions and requirements such as technological dependencies or resource availabilities have to be considered. This leads to a multitude of possible execution orders and consequently to a hard optimization problem. Within this paper the integration of a flexible Simulated Annealing metaheuristic into a constraint-based simulation concept is presented to determine optimized construction schedules. Simulated Annealing approximates the optimized solution by defined temperatures and cooling rates. Both parameters help to define the search process and respectively help to escape from local minima. In the following the Simulated Annealing approach and its implementation are described in detail. Furthermore, the application of Simulated Annealing is demonstrated using an example from construction scheduling.

Introduction

Scheduling the execution processes for a construction project is a challenging task. Resources such as machines and employees, as well as spatial restrictions of the construction site, must be considered in the scheduling of construction tasks. Additionally, the site layout changes during processing, which requires material and transport flows to be adapted. All-in, the execution flow is influenced by a multitude of different process and project requirements that have to be considered in detail during the planning phase. This leads to a multitude of execution sequences. Considering the project objectives, like the minimization of the total execution time or costs, an eligible alternative has to be generated and selected.

Simulation models are successfully applied within the manufacturing industry to improve production flows. The efficiency of production facilities, local plants or specific production lines can thereby be evaluated, and material flows as well as employee utilizations can be optimized. The application of simulation models is also a promising approach for planning various processes in the construction industry. Simulation enables users, for example, to visualise material flows, to localize manpower bottlenecks or to run what-if scenarios. Due to the fact that simulation applications in the manufacturing industry only support static layouts, a dynamic and flexible simulation approach is required to describe complex construction processes.

Within the research collaboration SIMoFIT (Simulation of Outfitting Processes in Shipbuilding and Civil Engineering) a constraint-based simulation approach was developed to improve execution scheduling of civil engineering and the shipbuilding industry (König et al. 2007). This approach allows, process and project conditions as well as current as-is states to be easily integrated by defining or removing constraints. During a simulation run the fulfillment of constraints is continuously checked and therefore only valid execution schedules are generated (Beißert et al. 2007).

The Simulated Annealing approach is integrated into the constraint-based simulation concept in order to calculate optimized schedules. Constituting on an initial schedule Simulated Annealing attempts to improve neighboring schedules by task substitution. Once an improved schedule is determined, the new solution replaces the current solution. Simulated Annealing enables declined solutions to be accepted in order to escape from local minima. The procedure and its consideration within the current simulation concept are presented within this paper in detail. Finally, a case study looking at the scheduling of several finishing trades is presented to evaluate the Simulated Annealing approach.

Constraint-based Simulation

Construction scheduling problems can be described by constraint satisfaction which is a powerful paradigm for modeling complex combinatorial problems (cf. Blazewicz et al. 2007). Classical constraint satisfaction problems are defined by sets of variables, domains and constraints between the variables (Rossi et al. 2006, Kumar 1992). Accordingly, when modeling the construction scheduling problems as constraint satisfaction problems, the construction tasks, material, employees, equipment, and the construction site layout are represented by variables. Subsequently, different scheduling constraints can be specified between these variables. These constraints have to be satisfied before the associated construction task can begin. Some typical construction constraints are formalized in Beißert et al. (2007) and shown in Table 1.

Table 1: Typical constraints for construction scheduling

Construction constraints
• Technological dependencies • Capacity, e.g. of equipment, employees • Availability, e.g. of material • Spatial aspects, e.g. safety or working areas

Variables and constraints can be represented by so-called constraint graphs. Thereby, for each constraint type an associated constraint graph can be calculated. Within this paper these graphs are used for the presentation of technological dependencies. The solutions of constraint satisfaction problems are valid execution orders of the construction tasks, where all associated technological constraints are fulfilled. Normally, solving such complex constraint satisfaction problems is extremely time-consuming. The constraint-based simulation can be used to generate a possible solution very quickly (cf. Beißert et al.). Therefore, the constraint satisfaction approach was integrated into a discrete event simulation application.

During discrete event simulations different events are generated by the procedures starting tasks and stopping tasks (cf. König et al.). Thus, the simulation time leaps from event to event. Furthermore the demand for the fulfillment of constraints is controlled within the procedures. A task can only be scheduled if all associated constraints are fulfilled. In Figure 1 the procedure of starting tasks is depicted. If a new event occurs, all tasks that have not been started are checked on fulfillment of their associated constraints. This leads to a set of next executable tasks. In the next step one of these executable tasks is selected for starting. Presupposed objects of this selected task like material, resources, or employees are locked during its execution and cannot be used by other tasks. Subsequently, all not-started tasks have to be checked again on fulfillment of their constraints by going to step one. The procedure is repeated until no more tasks can be started at the current time. If the remaining time of a construction task has expired, the task is marked as finished. Its presupposed objects are unlocked and can be used by other construction tasks.

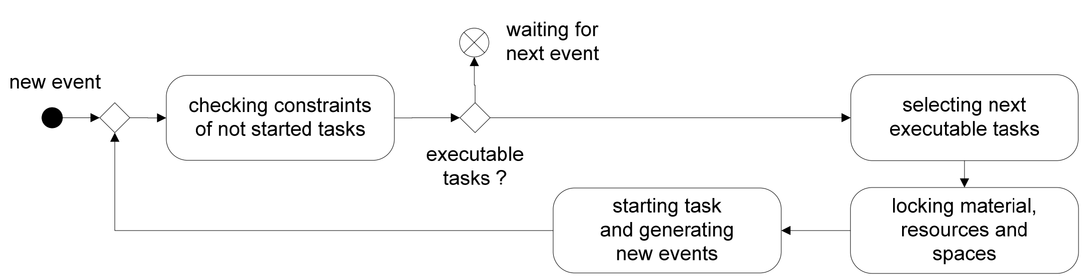


Figure 1: UML-diagram of starting tasks

The starting and stopping of construction tasks are repeated until all tasks have been completed. All events as well as the locking and unlocking of material, resources, equipment, and working spaces are recorded. Thus, each simulation run leads to a practicable construction schedule that can be investigated afterwards, for example with regard to time or costs.

Simulated Annealing

Simulated Annealing is a well-known local optimization approach for solving complex combinatorial problems like construction scheduling problems. The general goal of local optimization methods is to find good solutions in an adequate amount of time. The concept of Simulated Annealing is inspired by the physical annealing process in metallurgy (Kirkpatrick et al. 1983, Cerny 1985). In this context, physical annealing is known as the heating and controlled cooling of metal to bring the material structure from an arbitrary initial state to a state with the minimum possible energy. During heating, the metal atoms become unstuck from their current position and arrange themselves randomly. The slow cooling phase allows the atoms to find highly structured configurations with lower internal energy than in the initial configuration (cf. Aarts et al. 2005, Dréo et al. 2006).

If this physical process is considered as an analogy to our area of concern, the solutions of an optimization problem represent the possible configurations of the atoms. The objective value of a solution, the so-called cost factor, is equivalent to the internal energy state. Starting with a high temperature and a randomly selected initial solution, the Simulated Annealing heuristic calculates a new solution within the neighborhood of the current solution. The acceptance of new solutions is based on a probability that depends on the difference between the corresponding costs and on the current temperature. Ultimately this means, that high temperatures allow the acceptance of new solutions, which causes higher costs. The probability of accepting higher costs decreases within the optimization process. Once accepted, the new solution is the starting point for the next optimization step. Consequently, in order to use the Simulated Annealing heuristic an appropriate neighborhood, a good temperature-based probability and an effective decreasing rate for the temperature have to be specified.

Neighborhood

The definition of an appropriate neighborhood for a current scheduling solution is very important. The construction tasks and their associated technological constraints form a directed acyclic constraint graph. Topological ordering is used to calculate an execution rank for each task (cf. Pahl and Damrath 2000). The rank depends on the ancestor degree of the task and is determined by the maximum length of connected ancestors within the graph. Thus, tasks with the ancestor degree of k belong to the rank k. Consequently if all needed resources were available in unlimited quantities, all tasks with an identical rank could be executed simultaneously. Typically, however, construction tasks with the same rank can only be executed successively, depending on the resource capacities and the current resource allocations. This leads to a partial task order within each rank.

Within the scope of our Simulated Annealing approach the local neighborhood of a schedule is defined as the substitution of two construction tasks of the same rank. Thus, a new solution can be generated by selecting two different tasks of each rank for this substitution randomly. Figure 2 shows the topological ordering of the technological constraint graph of a simple scheduling problem. The problem consists of seven tasks {A, B, C, D, E, F, G} that have to be executed by four different resources {R1, R2, R3, R4}.

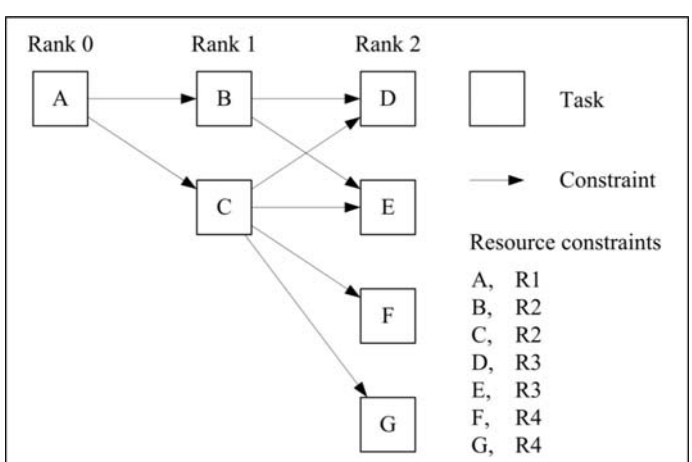


Figure 2: Topological ordering of construction tasks

The resource constraints are not considered when determining the topological ordering. Nonetheless, these constraints have a deep impact on the resulting schedules and therefore on the resulting makespan and project costs. An initial solution can be generated based on the specified resource requirements (cf. Figure 3). Within this solution each rank has also an initial execution order of its construction tasks. The initial execution order of rank one is < B, C > and of rank two is < D, F, E, G >.

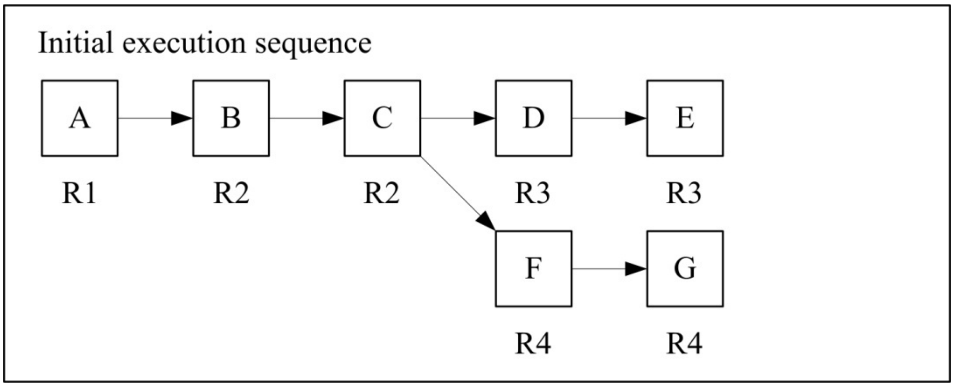


Figure 3: Initial execution sequence

Considering this initial solution a neighboring new solution can be generated by interchanging the execution positions of the tasks B and C for rank one. This new solution also fulfills the defined constraints and leads to another correct resource allocation. Now, the new partial order of rank one is {C, B} and of rank two is {E, F, D, G} (cf. Figure 4). In a second optimization step the execution positions of the tasks D and E belonging to rank two can be substituted.

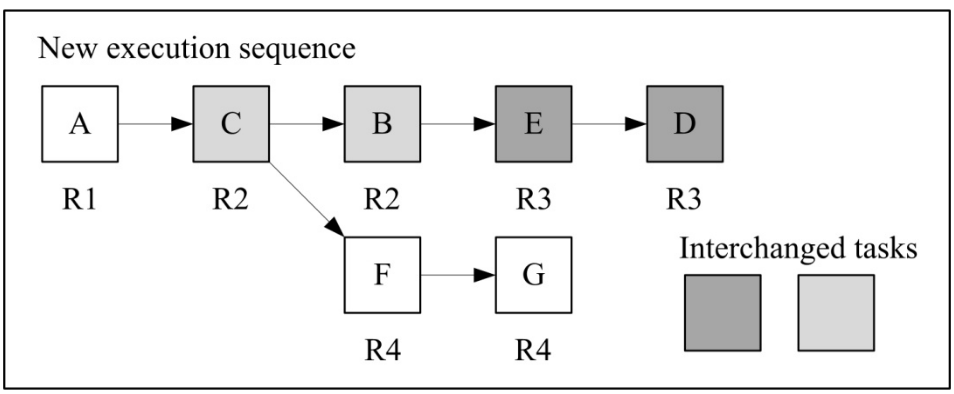


Figure 4: New neighboring execution sequence

Rules of acceptance

The probability of accepting a solution as a new start candidate can be specified by a probability function p(cc, cn, t), which depends on the current costs cc, the costs of the new solution cn and the current temperature t. Kirkpatrick et al. (1983) defined a typical probability function for many optimization problems as p(cc, cn, t) = 1 if cn < cc, and exp((cc – cn) / t) otherwise. This means that new solutions with lower costs are always accepted. When the temperature t goes to zero, the probability p tends towards zero.

The convergence speed of the Simulated Annealing approach to find good solutions depends on the initial temperature ti, the criterion for decreasing the temperature and the decreasing rate Δt of the temperature. Ideal values for these parameters cannot be determined beforehand. These control values strongly depend on the specific optimization problem and have to be empirically adjusted. Normally, the decreasing criterion is based on the length of homogeneous Markov chains or a number of maximal iterations for the same temperature (cf. Dréo et al. 2006). Different static and dynamic cooling concepts exist for decreasing the temperature (Aarts et al. 2005). Within our case study the decreasing criterion is defined by a maximum iteration number and the decreasing rate Δt of the current temperature t as Δt = t – (α * t) with the cooling parameter α. Typical values for α vary between 0.8 and 0.99 (Aarts et al. 2005). The detailed Simulated Annealing algorithm is shown in Figure 5.

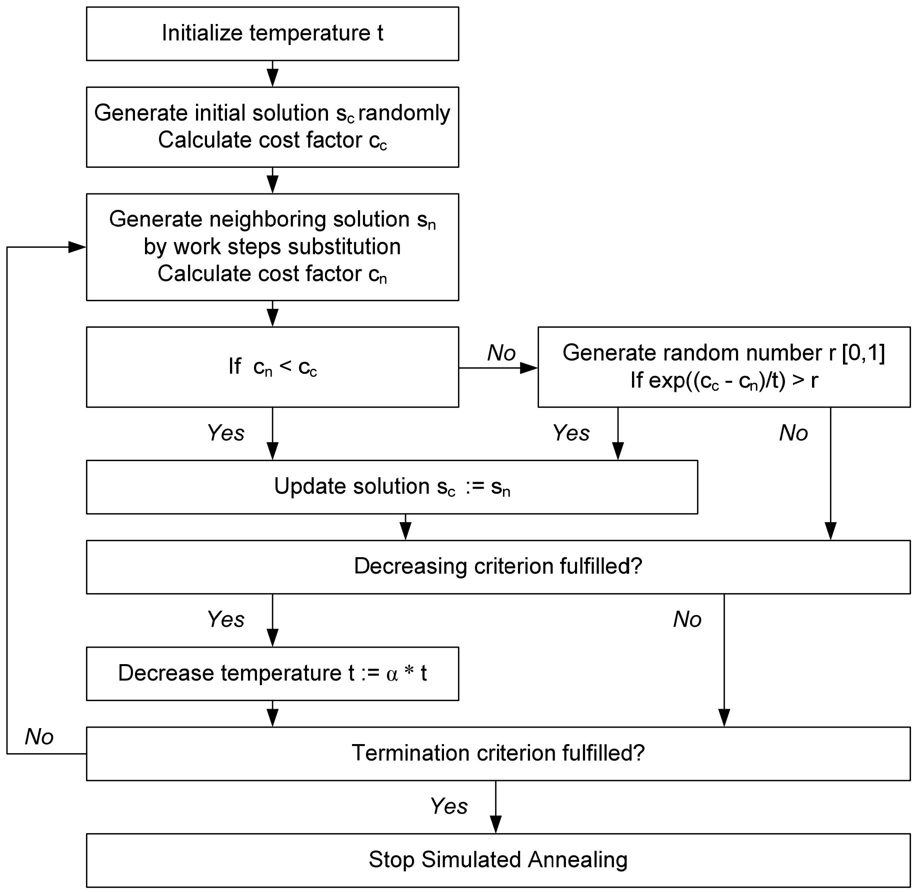


Figure 5: Simulated Annealing procedure according to Parmee (2001)

Implementation

Simulated Annealing is integrated into our constraint-based simulation approach by the following implementation steps. At the beginning of a new optimization the ranks of all construction tasks are generated, grouped, ordered, and stored in an execution list l. Furthermore the initial start temperature ti, the cooling parameter α, and maximum number of iterations for the same temperature are specified. For the generation of a new solution within a restricted neighborhood, the routine starting tasks (cf. Figure 1) was extended by the step "ordering tasks considering the restricted neighborhood" (cf. Figure 6). Therefore a temporary execution list lt is applied where the execution order of the tasks for each rank are stored. During a simulation run this list is continuously updated. After a list of next executable tasks is generated, its execution position within its associated rank is calculated for each task. Therefore, the temporary execution list is checked and the new results are added to the list. Thus, the execution sequence of each rank can be exactly specified afterwards.

Within the first simulation run the next executable tasks are started randomly. Based on the temporary execution list lt and the neighborhood definition further execution orders of the tasks are calculated for all subsequent simulation runs. New solutions are generated by random substitution of two different work steps of the same rank. The interchanging is started with the lowest ranks. If there are two or more next executable tasks listed belonging to the same rank, then two of them are interchanged. Within each optimization step only one rank modification is performed. Each generated execution order is stored in a Tabu list. If a new order already exists in the Tabu list, the ordering routine for the current event is repeated by interchanging other tasks iteratively until a new ordering is found or a termination criterion is fulfilled. The termination criterion is defined by a maximum number of iterations.

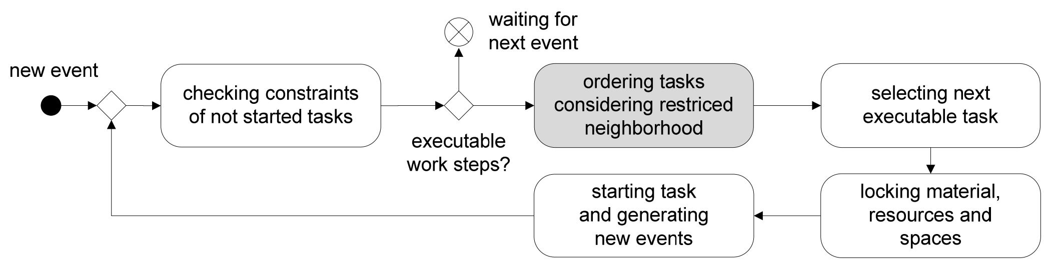
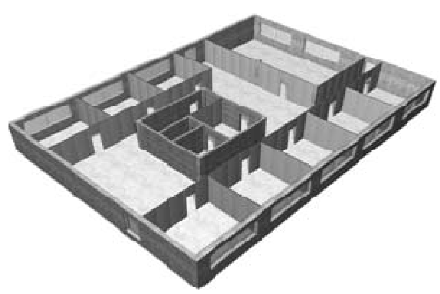


Figure 6: Generating a new solution by ordering executable tasks

After each optimization step the list lt contains a new execution order for all tasks. Consequently, the new costs cn for this new schedule can be calculated. In the next step, the probability value for accepting the new solution is specified and checked by using a normalized random number r. The new solution is accepted, if r is less or equal p(cc, cn, t), then the temporary execution list lt is copied to the list l. Following this, the current temperature t is reduced by Δt, if the maximum number of iterations for this temperature t is achieved.

Case study

In this section the presented Simulated Annealing concept is evaluated by scheduling several finishing trades of a building storey with twelve rooms, thirteen dry walls, twelve interior doors, and fourteen window sills (cf. Figure 7).



Building story (100 x 69 foot)

- 12 rooms
- 13 dry walls
- 12 interior doors
- 14 window sills

Figure 7: Finishing elements of building storey

The different finishing trades with their considered work steps are shown in Table 2. Based on the given number of rooms, dry walls, window sills, and interior doors as well as the specified finishing trades, 190 tasks and 1199 technological and resource constraints are generated.

Table 2: Finishing trades with tasks and required workers

Finishing Trades	*Tasks and required workers*
Ground works	spill floating screed (two floorer), lay flooring tile (one floorer), lay carpet (two floorer)
Dry construction	install drywall (two drywaller), install hung ceiling (two drywaller)
Joinery	install window sill (one joiner), install doors (one joiner)
Coating	color wall (one painter), color ceiling (one painter)

Figure 8 depicts the partial constraint graph for one floor, two hung ceilings, and two dry walls including the opening and coating tasks. In this example, the construction task *spill floating screed* has to be finished before covering tasks in the same floor can be started. Furthermore, before the construction tasks *spill balancing material* can be started the two associated walls have to be assembled completely.

For this case study two drywallers, two painters, two floorers as well as one joiner were specified. Under the given resource constraints (cf. table 2), optimized schedules with minimal makespan are determined by

Monte Carlo simulation and Simulated Annealing. Both optimization experiments were simulated for exact 60 minutes. Within the Monte Carlo experiment the next executable tasks are selected randomly. For the Simulated Annealing experiment an initial temperature ti = 300, a cooling parameter α = 0.8 and the same maximal number of iterations for the ordering routine (cf. section implementation) and for the temperature reduction of n = 10 were used.

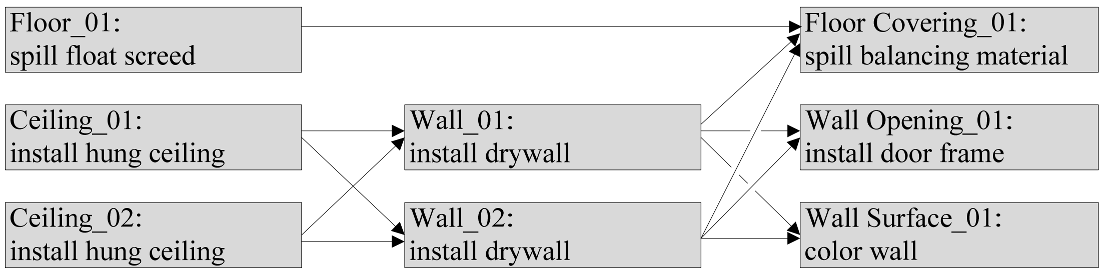


Figure 8: Partial constraint graph of selected finishing tasks

The results of the Monte Carlo and Simulated Annealing optimization are shown in Table 3. After 10 minutes 1000 different Monte Carlo schedules with an overall minimal makespan of 535 hours were simulated. By using Simulated Annealing 137 accepted solutions were generated. The last and therefore best solution has a makespan of 492 hours. In this case study the Simulated Annealing heuristic found an execution schedule that is about 43 hours faster than the best schedule generated by Monte Carlo simulation.

Table 3: Evaluation of the optimization experiments

Experiments (10 minutes)	Max [h]	Min [h]	$\bar{X}$ [h]	σ [h]
Monte Carlo	658	535	632	40.44
Simulated Annealing	627	492	525	19.63

Conclusions and Outlook

Using the constraint-based simulation approach different practical schedules can be simulated. Afterwards promising solutions can be evaluated in terms of work and material flow organization, utilization of space and worker's efficiency as well as process costs, afterwards. However, the application of Monte Carlo Simulation is very time-consuming and the locating of optimal respectively optimized schedules is not guaranteed. The application of local optimization methods guarantees the calculation of optimized schedules in an adequate amount of time. We integrate the Simulated Annealing metaheuristic in our simulation concept. Thus, clearly improved solutions can be generated in a diminished amount of simulation runs.

In future work, the justifications of the heuristic parameters like temperature and its cooling rate have to be investigated. The parameters choice is decisive for results quality. Thus, a problem-specific adaptation of both is essential to effective generate good schedules. The decreasing criterion based on the length of homogeneous Markov chains or a number of maximal iterations for the same temperature are promising possibilities for a problem-specific adaptation and are investigating in future work.

References

[1] Aarts, E., Korst, J., Michiels, W. (2005). "Simulated Annealing". In "Search Methodologies - Introductory Tutorials in Optimization and Decision Support Techniques" edited by Burk, E. K., and Kandall, G., Springer, New York

[2] Beißert, U., König, M., Bargstädt, H.-J. (2007). „Constraint-based simulation of outfitting processes in building engineering". 24th W78 Conference, Maribor, Slovenia

[3] Blazewicz, J., Ecker, K., Pesch, E., Schmidt, G., and Weglarz, J. (2007). „Handbook on scheduling: from theory to applications". Springer, Berlin

[4] Cerny, V. (1985). “A thermodynamical approach to the travelling salesman problem”. Journal of Optimisation Theory and Applications, Vol.15, 41-51
[5] Dréo, J., Siarry, P., Petrowski, A., Taillard, E. (2006). "Metaheuristics for Hard Optimization: Methods and Case Studies". Springer, Berlin
[6] Freuder, E.C., Wallace, R. (1992). “Partial constraint satisfaction”. Artificial Intelligence 58, 21–49
[7] Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P. (1983). “Optimization by Simulated Annealing”. Science, 220, 671-680
[8] König, M., Beißert, U., Steinhauer, D., Bargstädt, H.-J. (2007). “Constraint-based simulation of outfitting processes in ship building and civil engineering“. 6th Eurosim Congress in Modelling and Simulation, Ljubljana, Slovenia
[9] Kumar, V. (1992). “Algorithms for constraint satisfaction problems: A survey”. AI Magazine 13(1), 32-11
[10] v. Laarhoven, P.J.M., Aarts, E.H.L., Lenstra J.K. (1992). “Job shop scheduling by simulated annealing”. Operations Research 40(1), 113-12
[11] Pahl, P.J., Damrath, R. (2000). “Mathematische Grundlagen der Ingenieurinformatik”. Springer, Berlin
[12] Parmee, I.C. (2001). “Evolutionary and adaptive computing in engineering design”. Springer, London
[13] Rossi, F., van Beek, P., Walsh, T. (2006). “Handbook of Constraint Programming”. Elsevier, Amsterdam

Beyond Knowledge Management Platform: Design of Organizational Controls in Managing Knowledge

S. P. Ho[1], Y. Hsu[2] and E. Lin[3]

[1]Associate Professor, Department of Civil Engineering, National Taiwan University, Taipei, Taiwan 10617. e-mail: spingho@ntu.edu.tw,Webpage: http://140.112.10.60/spingho
[2] Assistant Professor, Department of International Business, National Taiwan University, Taipei, Taiwan 10617. e-mail: yhsu@management.ntu.edu.tw
[3]PhD Student, Department of Civil Engineering, National Taiwan University; e-mail: b92501100@ntu.edu.tw

Abstract

Knowledge has emerged or been regarded as the most strategically significant resource of the firm. From the resource-based view (RBV), the integration of individuals' specialized organizational capability through knowledge management (KM) has been considered crucial to the creation and sustainability of competitive advantages. Recently, KM has drawn immense attentions from researchers and practitioners in many industries and has become increasingly important in business management practice. However, most KM literature tends to focus on the "technology side" of KM such as platform design and knowledge warehousing issues. In practice, implementing KM platforms is also unfortunately considered the most important task of KM, if not the only task. As we've observed, many KM attempts died with their costly and advanced high-tech information and communication systems. In this paper, we argue that KM cannot be successful without appropriate organizational environment. For example, people in organizations are not well motivated to share knowledge and may even prefer not to share their knowledge in order to preserve their intellectual or proprietary values in organizations. This paper aims to study the behavioral dynamics of knowledge sharing in organizations and the design of organizational controls for KM from the perspective of knowledge sharing. First, a theoretic model based on game theory is developed to study the interaction between individuals' knowledge sharing behaviors and firms' strategies. Next, following the view of developed game model, a case study is conducted to study how knowledge is shared in different phases of KM implementation characterized by distinctive organizational controls. Finally, based on the game modeling, case analysis and critical study of literature, an integrated framework for the design of organizational controls in managing knowledge is proposed.

Keywords: Knowledge management, Knowledge sharing, Organizational controls, Game theory, Case study.

1. Introduction

Knowledge has emerged as the most strategically significant resource of the firm, and the integration of individuals' specialized organizational capability is crucial to the creation and sustainability of competitive advantages (Grant, 1996). From the resource-based view (RBV) that treats valuable resources as the cornerstones of competitive advantages, Knowledge Management (KM) can be considered one of the most critical processes for business success in today's intensively competitive environment. Recently, KM has drawn immense attentions from researchers and practitioners in many industries and has become increasingly important in business management practice. Surprisingly, there are only limited studies that address the most fundamental element in KM: the sharing of knowledge. In fact, people in organizations are not well motivated to share knowledge and may even prefer not to share their knowledge in order to preserve their intellectual or proprietary values in their organizations. Without the premise of each individual's willingness to share knowledge, there will be no valuable inputs for any KM platforms. Therefore, when organizations consider implementing costly Information and Communication Technology (ICT) platforms for KM, it is critical to assure that individuals are willing to share their knowledge through the platforms.

This paper aims to study the behavioral dynamics of knowledge sharing in organizations and the design of organizational controls for KM from the perspective of knowledge sharing. First, a theoretic model based on game theory is developed to solve for the conditions that determine the knowledge sharing behaviors of employees. Next, following the view of developed game model, a case study of STAR (pseudo name) Engineering Consultants Inc., one of the largest construction engineering firms in Taiwan, is conducted to study how knowledge is shared in different phases of KM implementation characterized by distinctive organizational controls. Finally, based on the game modeling and case analysis, an integrated framework for the design of organizational controls in managing knowledge is proposed.

2. Game Model of Knowledge Sharing

2.1 Game Theory

Game theory can be defined as the study of mathematical models of conflict and cooperation between intelligent rational decision-makers. There are two basic types of games: static games and dynamic games, in terms of the timing of decision making. In a static game, the players act simultaneously. On the contrary, in a dynamic game, the players act sequentially. Due to the nature of knowledge sharing, the dynamic game will be used for modeling here. Players in a dynamic game move sequentially instead of simultaneously. As to answer what each player will play/behave in this game, we shall introduce the concept of "Nash equilibrium," a set of actions that will be chosen by each player, and no player wants to deviate from the equilibrium solution. Thus, the equilibrium or solution is "strategically stable" or "self-enforcing" (Gibbons, 1992).

2.2 The Model and ITS Practical Implications

2.2.1 Parameters Regarding Employees

- γ_1: Explicit costs of sharing knowledge

The explicit costs exist because individuals have to invest time and efforts to share their knowledge. Therefore, one of the factors that affect individuals' sharing willingness is the explicit costs of sharing. The higher the costs are, the less the sharing willingness is.

- γ_2 : Implicit costs of sharing knowledge

When individuals share their knowledge, they may incur a hidden cost, due to the fact that their competitiveness and uniqueness in an organization may be hampered after sharing their specific knowledge. Such costs are conceptualized as the "implicit cost" of knowledge sharing. The magnitude of such costs depends on the uniqueness and importance of the knowledge within a particular organization.

- S: Intrinsic rewards of sharing knowledge

As individuals share their knowledge in organizations, they may get positive feedback from their colleagues or may enjoy better performance of their work groups. Sometimes one shares knowledge in order to gain respect or appreciation from colleagues or to build up professional reputation in a work group. These non-monetary rewards are regarded as "intrinsic rewards." In study, S is further broken down to S1, Reputation rewards, S2 , performance improvement of work teams, S3, altruism.

- ω: Extrinsic rewards for sharing knowledge

Extrinsic rewards such as monetary rewards are common means for promoting knowledge sharing.

2.2.2 Parameters Regarding Firms

It is assumed that firms will devise certain strategies for better knowledge sharing and management, such as organizational structures or incentive systems, characterized by the following parameters:

- c : Costs for ICT platforms and their implementation

The costs for ICT platforms and their implementation are usually accounted for the major costs of knowledge management and are often considered a costly investment.

- c_R : Firm's costs for providing extrinsic/monetary rewards

Under certain conditions as we shall discuss later, it would be beneficial for firms to provide monetary rewards for the sharing of particular knowledge. c_R will be one of the decisional factors for organizations in developing knowledge management strategy.

- π : Benefits due to improved firm competitiveness

The benefits due to better or improved competitive advantages are the major reason for KM. Although these benefits are difficult to be precisely quantified, we assume that in practice they can be estimated so as to compare with the costs for KM.

2.2.3 Game Modeling and Knowledge Taxonomy

We assume that knowledge sharing game is a dynamic game with complete information, where firms first provide the environment for knowledge sharing, and then employees decide to share or not. Figure 1 shows the game model of knowledge sharing dynamics. As shown, there are two players in the game, the employee and the firm. The strategy of the firm and the sharing decisions of the employees are made based on the characteristics of the knowledge in the firm. By solving the game trees in Fig. 1 backward recursively, we obtain five possible game equilibria and seven corresponding scenarios numbered in Fig. 1. Due to the length limit, the detailed derivations of each equilibrium and scenario can be found in Ho et al. (2006).

Six types of knowledge are identified or categorized as shown in Table 1 based on the equilibria characteristics. The six types of knowledge are characterized by three dimensions: 1. explicit costs of sharing knowledge, γ_1 , 2. implicit costs of sharing knowledge, γ_2 , and 3. benefits to the firm due to the sharing, π . First, the "simple knowledge" is characterized by low explicit and implicit sharing costs, as characterized by scenarios 1 and 4. A simple database system or sharing community platform would be sufficient for managing scale-sensitive simple knowledge. Second, the "core simple" knowledge is the simple knowledge with the scale economies in dissemination and utilization of the knowledge. Third, the "core non-unique knowledge" is characterized by high explicit but low implicit sharing costs, and high benefits to the firm. Firms need to either largely reduce the explicit sharing cost or provide rewards in order to encourage the sharing. Fourth, the "core unique" knowledge is characterized by much higher implicit sharing costs compared to the core non-unique knowledge. Firms usually need to provide extrinsic rewards for encouraging the sharing. Fifth, the "special knowledge" is characterized by low benefits to the firm compared to the core unique knowledge. For example, knowledge on a firm's IT infrastructure or on tax laws can be regarded as special knowledge. According to the equilibrium, extrinsic rewards are needed to encourage the sharing, but it is not economical to provide such incentives. Sixth, the "spurious knowledge" is characterized by high explicit sharing but low benefits to the firm. The sharing of such knowledge, such as out-dated or irrelevant knowledge, should not be encouraged.

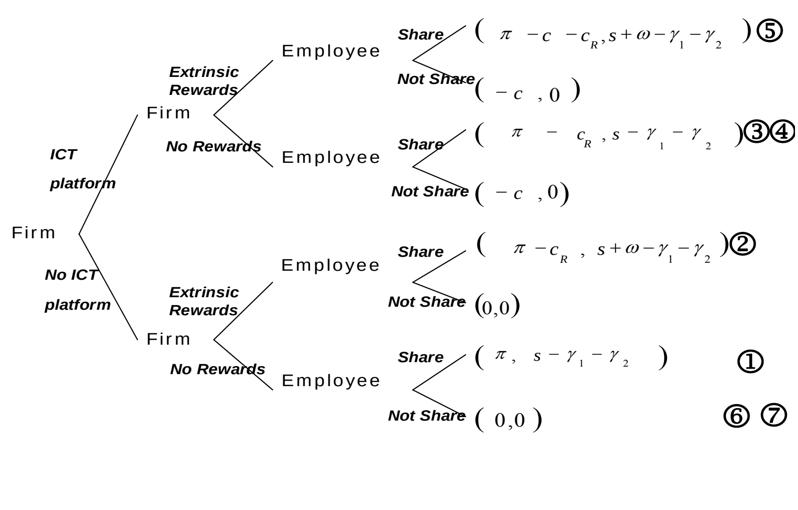


Figure 1. Game Model of Knowledge Sharing Dynamics

Table 1. Knowledge Categorization and Equilibrium Scenarios

	π (L)		π (H)	
	γ_2(L)	γ_2(H)	γ_2(L)	γ_2(H)
γ_1(L)	Simple knowledge ①	Special knowledge ⑦	Core simple knowledge ④	Core Unique knowledge ② ⑤
γ_1(H)	Spurious knowledge ⑥		Core non-unique knowledge ③	

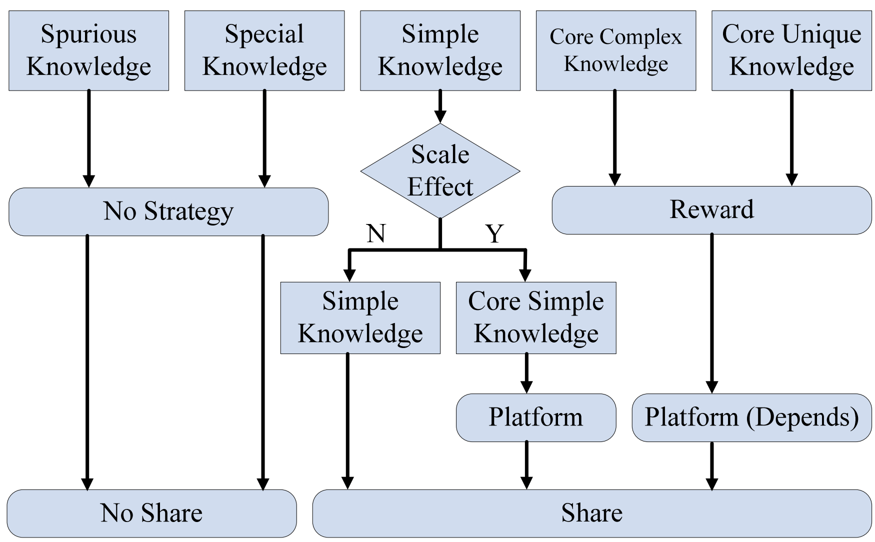


Figure 2. Model Implied Strategies for Knowledge Sharing

3. Case Study

3.1 Background

STAR (pseudo name) Engineering Consultants Inc. is one of the largest engineering consulting firms in Taiwan, also one of the ENR top engineering consulting firms in the world. Beginning from the year 2000, Knowledge Management System (KMS) was implemented to promote its working efficiency and maintain its competitiveness. STAR employed a system that quantifies the sharing efforts of each individual and aims to ensure the minimum level or points of sharing of each individual. Two major ICT platforms were implemented: the KM database system and the knowledge sharing and communication platform, including a so called "Emergency" bulletin (pseudo bulletin name) for timely problem solving and discussions. Several studies, such as Bartol and Srivastava (2002), Cabrera and Cabrera (2002), Pascarelia (1997), suggest that a useful approach for tacit knowledge sharing is to develop various internet-based communities where employees and experts with specific knowledge or skills can gather together to share their knowledge through discussion.

3.2 Survey Study of the Case

3.2.1 Survey Design

The total number of questionnaires sent out was 1164. 958 questionnaires came back and 806 of them were valid for analysis. The survey was designed to verify the hypothesis concerning the knowledge sharing model. Regarding the knowledge type characterized by the aforementioned three dimensions: γ_1, γ_2, and π, we used three questions to evaluate the three characteristics of the knowledge owned by individuals. Six-point Likert scales were used in the questionnaire to prevent neutral answers to the questions, where scale one was strongly disagree (or very low) and six was strongly agree (or very high). In order to categorize different types of knowledge using the measurements of γ_1, γ_2, and π, we considered Likert scales one to three as low values and scales four to six as high values.

3.2.2 Survey Results and Implications on Knowledge Sharing

- Knowledge distribution and the sharing willingness

As shown in Table 2, according to the proposed categorization the individuals who own core unique, core non-unique or core simple knowledge are about 67% of the total respondents. Given that STAR is an engineering consulting firm, the knowledge profile and distribution as shown in Table 2 is considered quite reasonable according to a senior manager of STAR. Table 2 also shows the statistics of individuals' willingness to share and perceived payoffs from sharing with respect to the types of their knowledge. The ANOVA p-values in each column of Table 2 are less than 0.001, indicating that the distinction between different categories is statistically significant. According to the survey, the individuals who own core-type knowledge have higher willingness of sharing than those who own spurious and special knowledge. This indicates that the knowledge sharing incentive system of STAR was in the right direction in terms of the types of knowledge to be encouraged for sharing. However, the mean willingness to share core (unique and non-unique) knowledge is still far from scale five or six, showing that although many efforts were employed to encourage the sharing of knowledge, these efforts were not very effective. We shall discuss later what could be more effective means to promote the sharing of core knowledge.

- Perceived payoffs due to knowledge sharing

Table 2 also shows four major possible payoffs from sharing knowledge perceived by different types of knowledge owners. These payoffs include the extrinsic rewards, omega ω, and the intrinsic rewards, S. According to Table 2, the extrinsic rewards are significantly lower than others in all types of knowledge. The highest and lowest payoffs among intrinsic rewards are altruism and reputation, respectively. We shall discuss their impacts on knowledge sharing later in details.

Table 2. Sharing Willingness and Perceived Payoffs

Knowledge Types	%	Sharing Willingness	Omega	S1	S2	S3
ALL	100 %	**3.484**	2.548	3.553	3.725	4.145
Spurious Knowledge	12.3%	**3.051**	2.091	2.929	2.939	3.778
Special Knowledge	5.5%	**3.295**	2.477	3.523	3.614	3.932
Simple Knowledge	14.8%	**3.025**	2.378	2.697	2.908	3.496
Core Simple Knowledge	12.0%	**3.825**	2.526	3.835	3.990	4.079
Core Non-unique Knowledge	19.6%	**3.573**	2.373	3.696	4.013	4.316
Core Unique Knowledge	35.8%	**3.689**	2.889	3.952	4.100	4.388

- Who are the owners of different types of knowledge?

Table 3 shows the demographic profile of different types of knowledge, which may help us identify the owners of each type of knowledge, as shown in the last column of the table. For example, we consider that the owners of simple knowledge are mostly experienced supporting staff because its demographic profile shows the highest percentage of female employees, lowest education level and relatively low salaries. Particularly, we find that the core unique knowledge is mainly owned by senior engineers, characterized by highest level of education, relatively high pay and medium seniority. The ANOVA p-values in each column of Table 3 are less than 0.001. When our conclusion concerning knowledge owners was presented to STAR's top managers, they considered such matching between knowledge owners and specific knowledge reasonable. The purpose of locating the owners of different types of knowledge is twofold. First, a reasonable match between knowledge owners and knowledge indicates that the knowledge taxonomy implied by the knowledge sharing model was supported by the data. Second, characteristics of the owners of specific knowledge may help to devise further strategies for sharing knowledge. For example, since the senior middle to high level managers typically do not have time to sit in front of computers for an extended amount of time, to promote their sharing of the core non-unique knowledge, it is crucial to include other sharing channels or approaches that are not computer based, such as mentoring.

Table 3. Knowledge Owners and Their Demographic Profiles

Knowledge types	Age	Salary (K)	Education	Years	Male%	Who are the Owners?
Spurious (99)	37.55	49.29	4.40	8.68	62.6%	Junior engineers and supporting staff
Special (44)	39.48	52.39	4.25	10.55	72.7%	Professional supporting staff
Simple (119)	39.34	50.29	4.04	11.05	58.0%	Experienced supporting staff
Core Simple (97)	43.10	64.27	4.30	13.71	81.4%	Senior lower level managers
Core Non-unique(158)	43.60	69.56	4.51	13.78	79.1%	Senior middle to high level managers
Core Unique(289)	41.39	61.21	4.52	11.98	76.8%	Senior engineers

Salary: monthly pay, in Taiwan Dollars; Education: 3=senior high, 4=Bachelor's, 5=Master's; Years: number of years working in this firm.

4. Implications on Organizational Controls for Better Knowledge Sharing

4.1 *Focusing on the Sharing of Core Types of Knowledge*

- Only core types of knowledge deserve specific resources for promoting the sharing

Since only core types of knowledge can bring enough benefits to cover the costs of resources for managing knowledge, the organizational arrangements or strategies for knowledge sharing should make sure that core types of knowledge are shared. A good criterion to identify core knowledge is to evaluate whether the sharing of the knowledge can enhance an organization's competitive advantage.

- The sharing of non-core types of knowledge will pollute the knowledge bases and hamper the utilization of KM systems

Garbage In Garbage Out (GAGO) also applies to KM systems. The more useless or spurious knowledge is in the systems, the less the individuals of an organization is likely to benefit from or utilize the systems. As a result of oversharing of non-core types of knowledge, either the individuals are less motivated to use the systems or the organization has to incur significant costs to differentiate the quality of shared knowledge.

- Use different "individual's minimum sharing requirements" for different types of knowledge owners

The use of "individual's minimum sharing requirements" may help to ensure the overall minimum level of sharing that may help to alleviate the social dilemma complex (Cabrera and Cabrera, 2002); i.e., why I am the only one contributing. However, according to our case study, the use of uniform sharing requirements may force those own non-core types of knowledge share low value knowledge or even garbage in order to meet the requirements. An organization should try to identify the owners of different types of knowledge similar to those shown in Table 3 and impose lower requirements for the owners of non-core types knowledge, such as the new comers or supporting staffs.

4.2 *The Use of Extrinsic Rewards*

- Using extrinsic rewards but mainly as the symbol for reputation

Although our game model indicates that extrinsic rewards are one of the major incentives that an organization can use to promote the sharing of core unique and non-unique knowledge, the case study shows that the use of high rewards may not be a good strategy because 1. high extrinsic rewards usually encourage high frequency of sharing, particularly the sharing of easier-to-share non-core types of knowledge and 2. it is too costly to differentiate the quality of shared knowledge. However, extrinsic rewards can be an excellent symbol for reputation of sharing core types of knowledge; especially the rewards are given by the high ranked manager in an open setting.

4.3 *The Use of Intrinsic Rewards*

- The intrinsic rewards can be critical incentives for the sharing of core types of knowledge

As shown in Table 2, if the extrinsic rewards cannot be too large as argued above, then the intrinsic rewards must be high enough to promote the sharing of core types of knowledge. Reputation is usually associated with the quality of the shared knowledge (how many people find it helpful). Performance improvement of work teams is often achieved when project/department leaders or managers share their knowledge. Altruism is found to be the highest among all perceived payoffs in the studied case. This indicates that the culture of helping each other, cooperation, or team spirit will strengthen the altruism rewards and hence enhance the sharing willingness.

4.4 *The Use of ICT Platforms*

- Advanced ICT platforms should be used only when the platforms can help to achieve either scale economy in disseminating the knowledge or significantly lower explicit costs of sharing.

The use of advanced ICT platforms does not necessarily increase the probability of a successful KM. For example, for organizations of small scale or non knowledge-intensive, basic ICT platforms may be enough for managing knowledge.

4.5 *Three Stages of Promoting Knowledge Sharing*

- First stage: using carrot and stick to break an old culture

Carrot and stick strategy is used when the systems are implemented. Carrot is used to promote sharing. Over-sharing and the sharing of non-core types of knowledge should not be a concern. The purpose of sharing at this stage is to familiarize the individuals with the systems and to establish the habits of sharing. Stick is used to maintain a minimum level of sharing within an organization to reduce the social dilemma problems discussed previously and also to help individuals establish the habits of sharing through systems.

- Second stage: shaping a new culture

In this stage, extrinsic rewards are reduced and intrinsic rewards are emphasized. Organizations begin to focus on the sharing of core-types of knowledge and shaping the culture that strengthen both the magnitude and the effect of intrinsic awards.

- Third stage: sharing core types of knowledge as a culture and an institution

In this stage, organizations mainly focus on using intrinsic rewards. Punishments for not meeting the minimum sharing level are significantly reduced.

5. Conclusions

In this study we introduce a game-theoretic model for analyzing the knowledge sharing dynamics and defined six types of knowledge, characterized by three dimensions. In our corporate-wide single case study, we find that the survey results are largely consistent with model implications. Combing the model and case study, we derive organizational control strategies and propose some illustrative practices for better managing knowledge.

Acknowledgement

The support of STAR (pseudo name) consulting firm is greatly appreciated. Without their support, this study is not possible.

References

[1] Bartol, K.M. and Srivastava, A. (2002). "Encouraging Knowledge: The Role of Organizational Reward Systems," Journal of Leadership and Organization Studies, 9 (1).
[2] Cabrera, A. and Cabrera, E.F., (2002). "Knowledge-Sharing Dilemmas," Organization Studies, 23 (5), 2002.
[3] Gibbons, R. (1992). Game Theory for Applied Economists. Princeton University Press, Princeton, N.J.
[4] Grant, R.M. (1996). "Prospering in Dynamically Competitive Environments: Organizational Capability as Knowledge Integration," Organization Science, 7 (4), pp. 375-387.
[5] Ho, S. P., Wu, P., and Hsu, Y. (2006). "Knowledge Sharing Model and its Implication on Knowledge Categorization and Management." Proceedings of the 23rd ISARC, Tokyo, Japan.
[6] Pascarella, P. (1997). "Harnessing knowledge," Management Review, 86 (9): 37-40.

Applying Neural Network Ensemble Concepts for Modelling Project Success

Yu-Ren Wang[1], Yi-Jao Chen[2] and C.F. Jeffrey Huang[3]

[1]Assistant Professor, Department of Civil Engineering, National Kaohsiung University of Applied Sciences, 415 Chine-Kung Road, Kaohsiung Taiwan, 807; PH (886) 73814526-5250; FAX (886) 73831371; email: yrwang@cc.kuas.edu.tw

[2]Lecturer, Deptartment of Asset Science, Leader University, No. 188, Sec. 5, An-Chung Rd., Tainan 70946, Taiwan, PH(886)62558019; FAX(886)62558528; email: yjchen@mail.leader.edu.tw

[3]Associate Professor, Department of Civil Engineering, National Kaohsiung University of Applied Sciences, 415 Chine-Kung Road, Kaohsiung Taiwan, 807; PH (886) 73814526-5234; FAX (886) 73831371; email: jeffrey@cc.kuas.edu.tw

Abstract

Researchers have been studying early planning process since the early 1990's and results from these researches suggest that projects with more early planning efforts are more likely to succeed. This study intends to employ neural networks to build credible models linking preproject planning and project success. Preproject planning status as measured by Project Definition Rating Index (PDRI) is set as the independent variable and schedule/cost performance is set as dependent variable. To enhance the performance of the neural networks model, bootstrap aggregation and boosting algorithms are incorporated in the model development process. The results from these two neural network ensemble models are examined. This research finds out that boosting neural network ensemble models produce better results than bootstrap aggregation neural network ensemble models. Results from both models show that project with better early planning can expect better chance of project success.

Keywords: Preproject Planning, Project Success, ANNs Model

1. Introduction

It has long been recognized by the industry practitioners that efforts made early in the project life cycle will have significant impact on the project outcome. Nevertheless, the early planning process varies significantly throughout industry from one organization to another. In light of this, researchers at the Construction Industry Institute (CII) and University of Texas at Austin have set up several research projects focusing on the early planning process since the 1900's. In the early planning stage of the project life cycle, the process of developing sufficient strategic information with which owners can address risk and decide to commit resources to maximize the chance for a successful project is defined as the pre-project planning process (CII 1995). Preproject planning is a major phase of the project life cycle and it begins after a decision is made by the business unit to proceed with a project concept and continues until the detailed design is developed.

Preproject planning process constitutes a comprehensive framework for detailed project planning and includes scope definition. Project scope definition is the process by which projects are selected defined and prepared for definition. It is a key practice necessary for achieving excellent project performance (Merrow and Yarossi 1994) and is a key element in the preproject planning process. How well preproject planning is performed will affect cost and schedule performance, operating characteristics of the facility, as well as the overall financial success of the project (Gibson and Hamilton 1994). Success during the detailed design, construction, and start-up phases of a project highly depends on the level of effort expended during the scope definition phase as well as the integrity of project definition package (Gibson and Dumont 1996). Therefore, it is important to investigate the relationship between preproject planning and project success with real data from the industry.

In order to measure the preproject planning efforts for each construction project, a scope definition tool, Project Definition Rating Index (PDRI) is incorporated in this research to evaluate the completeness of project scope definition. The Project Definition Rating Index, developed by CII, is a comprehensive, weighted checklist of crucial scope definition elements that have to be addressed in pre-project planning process. It provides the project team a simple and easy-to-use tool to objectively evaluate the current status of a project during preproject planning phase. Since its development, researchers at the University of Texas at Austin and Construction Industry Institute (CII) have been collecting preproject planning information using the PDRI. For the uniqueness of the different sectors in the construction industry, two versions of the PDRI have been developed specifically for the Industrial and Building sectors.

In addition to preproject planning information collected using the PDRI, project performance (cost and schedule) information was also collected through the data collection process. Enhanced Artificial Neural Networks (ANNs) models, using bootstrap aggregation and boosting algorithms, are selected in this research to investigate the relationship between preproject planning and project performance using the sample project data. The ANNs analysis results from the two different models are also compared.

2. Survey Instrument and Data Collection

A scope definition tool, Project Definition Rating Index (PDRI) is used as a survey instrument in these case studies to measure the preproject planning practices in the industry. The PDRI is developed by CII in 1994 as an easy-to-use preproject planning tool to assist owner and contractor companies to better achieve business, operational, and project objectives (CII 1996). The first version is the PDRI for industrial projects, which is a weighted matrix with 70 scope definition elements (issues that need to be addressed in preproject planning) grouped into 15 categories and further grouped into three main sections. Since its development, the Industrial PDRI has been widely used among the CII member companies. In responding to the needs of the building industry, CII chartered a research team and developed the PDRI for Building Projects in 1999 (CII 1999).

The PDRI provides a means for an individual or team to evaluate the status of a construction project during preproject planning with a score corresponding to the project's overall level of definition. The PDRI helps the stakeholders of a project to quickly analyze the scope definition package and to predict factors that may impact project risk specifically with regard to industrial and building projects (Cho 2000).

In the PDRI survey questionnaires, specific questions were intended to obtain historical and "after the fact" project information. The questionnaires included questions regarding project basics (location, type, budget and schedule), operating information, and evaluation using an unweighted PDRI score sheet. Survey participants were asked to think back at a point just prior to construction document (detailed design) development when they filled out the PDRI evaluation score sheet. The total scores were then calculated based on pre-assigned element weights after the questionnaires were returned. Please refer to CII 1996 and CII 1999 for detailed development of PDRI element weights. Due to the unique nature of these two different sectors, industrial and building projects were examined separately throughout this research investigation.

The data collection was accomplished through a series of retrospective case studies. The sample projects used in this study were obtained from three different sources: previous PDRI research, CII Benchmarking and Metrics research, and institutional organizational (which prefers remaining anonymous) PDRI benchmarking research. Data from 62 industrial projects and 78 building projects, representing approximately $5 billion in total construction cost, were collected and used to conduct an investigation of the early planning practices in the industrial and building industry. Nevertheless, it is important to note that the collected sample from these three sources is based on organization's volunteering projects and not on a random sample of a known population.

3. Data Analysis and Modeling

Information related to preproject planning practice and project performance is collected through the sample project survey. A database is set up for this research project and is used for data storage and further analysis. The preproject planning status for each surveyed project is measured using the PDRI evaluation. Two project performance aspects are of particular concern for this research: cost and schedule performance. Cost performance and schedule performance are measured by cost and schedule growth. In the survey, respondents were asked to provide estimated costs at the start of construction document development as well as the actual costs after construction completion. Total cost growth measures total project cost growth as a percentage of the initial estimated project cost. Cost performance was measured by project Cost Growth metric obtained as follow:

$$\frac{\textit{Final Cost} - \textit{Initial Estimated Cost}}{\textit{Initial Estimated Cost}} \quad (1)$$

The total project duration used to calculate project schedule growth was measured from the start date of construction documents development to the date of substantial completion in months. The following equation was used for computing project schedule performance, Schedule Growth:

$$\frac{\textit{Final Schedule} - \textit{Initial Estimated Schedule}}{\textit{Initial Estimated Schedule}} \quad (2)$$

When considering cost performance, it is defined in this research that if the project has a cost growth equal to or less than 0, the project is classified as a successful project. In the mean time, when considering schedule performance, a project is defined as a successful project if the schedule growth is equal to or less than 0. It should be noted that schedule and cost performances are investigated separately throughout this research.

Two different algorithms, bootstrap aggregation and boosting, are used to develop Artificial Neural Networks models for this research to investigate the relationship between the preproject planning, as measured by PDRI scores, and project success, as measured by cost/schedule growth.

4. Artificial Neural Networks

The principle of Neural Networks is based on the assumption that a highly interconnected system of simple processing elements can learn complex interrelationships between independent and dependent variables (Elhag 2004). A typical neural network consists of an input layer, an output layer, and one or more hidden layers. These layers are connected by neurons to form a parallel distributed processing system. Each neuron is viewed as a processing element (PE) that receives inputs and generates outputs through an activation function. Each of the connections between the process elements has an associated weight. Figure 1 shows a typical three-layered neural network with an input layer (I), a hidden layer (H), and an output layer (O).

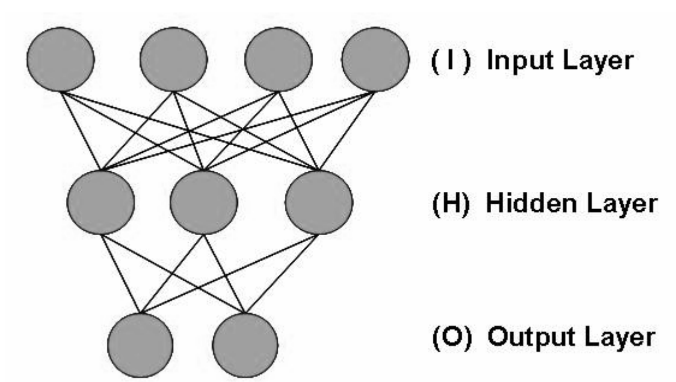


Fig. 1. An example of a three-layered neural network

In the hidden layer, each neuron receives an activation signal (input), and generates a signal (output) through an activation function. The activation signal is the weighted sum of all the signals entering the neuron, as shown in Eq. (3). In Eq. (3), xj is the activation signal that the neuron j in the hidden layer receives; Ii is ith input in the input layer; and Wij is the weight of the connection between the neuron j in the hidden layer and the input Ii. The neuron (Process Element) produces an output through an activation function that can be any form. The activation function can be either linear or non-linear, and one of the most commonly used activation function is the sigmoid function. The general form of sigmoid function is shown in Eq. (4), where hj equals output of the neuron j in the hidden layer and xj equals input for the neuron j.

$$x_j = \sum_i I_i W_{ij} \tag{3}$$

$$h_j = f(x_j) = \frac{1}{1+e^{-x_j}} \tag{4}$$

$$y_k = \sum_j h_j W_{jk} \tag{5}$$

As presented by Eq. (5), the neurons in the output layer receive activation signals (weighted sum of inputs to neuron k) from the neurons in the hidden layer. In Eq (5), yk is the input of the neuron k in the output layer and Wjk is the weight of the connection between the neurons j and k in the hidden and output layers, respectively. In the output layer, these activation signals are transformed (through activation function) again to generate the outputs of the neural network. This process is shown in Eq. (6), where ok is the predicted value of the outputs. Then the outputs are compared with desired or actual values, dk.. The error (difference between predicted value and desired/actual value) at the output neurons is defined by Eq. (7). The best performance of the neural network is achieved when the error is minimized.

$$o_k = f(y_k) = \frac{1}{1+e^{-y_k}} \tag{6}$$

$$E(W) = \frac{1}{2}\sum_k (d_k - o_k)^2 \tag{7}$$

For supervised neural networks (models with specific actual/desired outputs), one of the most effective and popular technique to minimize the error function E(W) is the back-propagation (BP) algorithm. For back-propagation neural networks, the error at the output layer propagates backward to the hidden layer and then to the input layer to update the weights for each of the connections in the neural networks. These forward process (input layer to hidden layer to output layer) and backward process (output layer to hidden layer to input layer) are repeated to minimize the error.

These repeated processes are viewed as learning (training) process. The relationships between inputs and outputs of the system are memorized through the connection weights. It should be noted that before the learning process starts, small random numbers (e.g., between −0.1 and 0.1) are assigned as the initial weights to the connections between the neurons. This ensures that the network is not saturated by large values of the weight, and prevents some training pathologies. Sometimes, the data will be normalized before to obtain convergence within a reasonable number of cycles.

Studies have shown that ANNs have sevreal advantages over the traditional statistical methods such as multiple regression analysis and multivariate analysis (Elhag and Boussabaine 2002). ANNs does not require that the data must follow a specific statistical distribution and does not require predetermination of the relationships between inputs and outputs. In addition, ANNs have very strong capability of self-learning and self-updating. Despite the advantages above, ANNs has been criticized as "unstable algorithm" (Roiger and Geatz 2003) because the model is very sensitive to slight changes in the training data. In view of this, Hansen and Salamon (1990) proposed the concept of neural network ensemble. Instead of developing one neural work to solve a particular problem, a group of neural networks are developed and the network output for each of them are then integrated to obtain a solution for the problem. Bootstrap aggregation and boosting algorithms are two applications of neural network ensembles that can enhance the performance of neural network models. The concept of neural network ensemble (compared with traditional method) is shown in Figure 2 below.

Breiman (1996) used bootstrap method to generate several sub-datasets from the original training data. These sub-datasets are fed into classification tree, composed of a group of classifiers, for training, and the final model output is obtained through the aggregation of the group classifier outputs. It is found bootstrap aggregation is able to reduce the model error rate by 20% to 47% when comparing with single classifier model. For this research, neural networks are created to serve as the classifiers in the bootstrap aggregation (or bagging) method. The outputs from the neural network classifiers are integrated to produce the final output of the model. As illustrated in Figure 3, bootstrap aggregation method is incorporated in the neural network ensemble for this research.

Freud and Schpire (1996) applied boosting algorithms obtained a better result when comparing with bootstrap aggregation method. Similar to bootstrap aggregation method, boosting method also creates a group of classifiers through a series of training process. However, the selection of training data is different. For bootstrap aggregation method, the training dataset, also known as bootstrap sample, is chosen at random with replacement from the original training set. Boosting maintains a weight for each instance and the higher the weight, the more likely the instance will be selected for classifier training. At each training process, the weight of misclassified instance is increased and thus makes it more likely to be chosen for the next round. After a series of classifier training under boosting approach, the final classifier aggregates the learned classifiers and the results are better than bootstrap aggregation method. Using neural networks as classifiers, the boosting approach is shown in Figure 4.

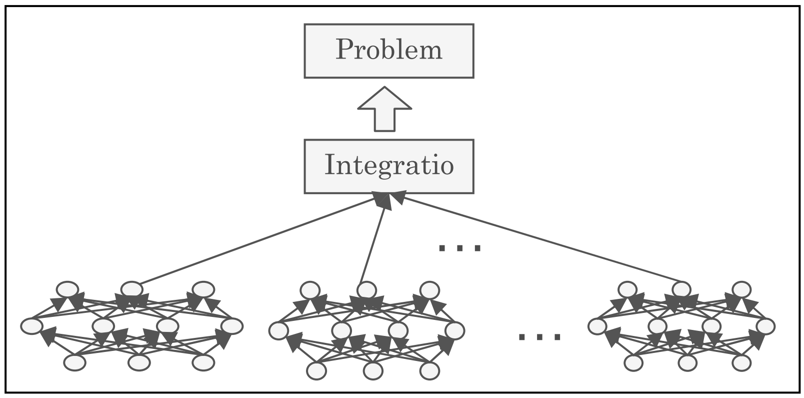


Fig. 2. Neural Network Ensemble

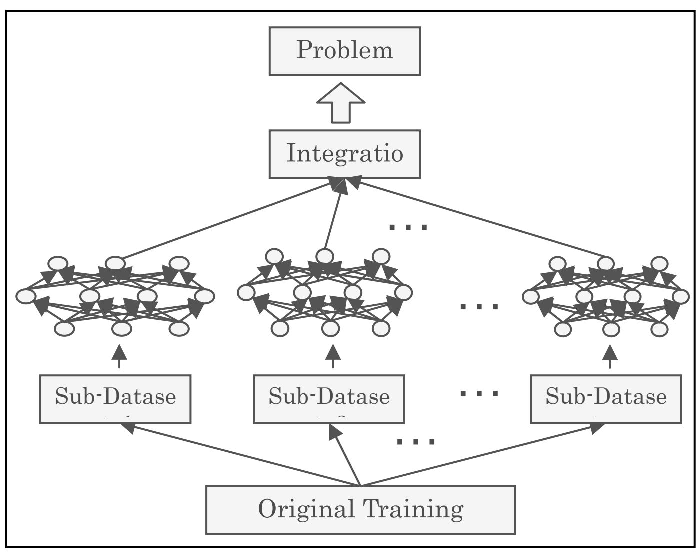


Fig. 3. Bootstrap Aggregation Neural Network Ensemble

Information from a total of 62 industrial and 78 building projects is collected and used to investigate the relationship between the preproject planning and project success. Preproject planning status, as measured by the PDRI score, is the independent variable while the project success, as measured by cost/schedule growth, is the dependent variable in this research. Bootstrap aggregation and boosting algorithms are incorporated to create neural network classifiers using the PDRI scores as model inputs. The sample projects are classified as successful or less-than-successful projects based on their cost and schedule performances.

There are three neural network models developed for this research: single classifier neural network, bootstrap aggregation neural network ensemble model and boosting neural network ensemble model. Single classifier neural network is the traditional neural network model development approach, which uses one neural network classifier for prediction. In this research, the average error rate (misclassification) for single classifier neural network model is 15.79%. Bootstrap aggregation neural network ensemble model trains a group of classifiers and the final model prediction is obtained by voting from the group outputs. The average error rate is 10.52% under this model. The average error rate obtained for boosting neural network ensemble model is 7.89%.

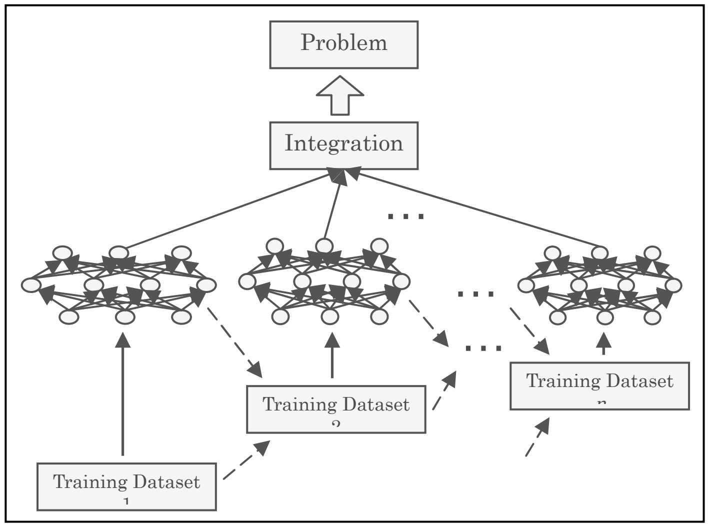


Fig. 4. Boosting Neural Network Ensemble

5. Conclusions

This paper studied the preproject planning of industrial and building construction projects and investigates its relationship with project success (as measured by cost and schedule performance). Questionnaire surveys were used to obtain information related to the status of preproject planning and project performance. Based on the collected data, this research developed three neural network models to classify the projects as successful and less-than-successful projects using the PDRI scores. The first model is traditional single classifier neural network model; the second is bootstrap aggregation neural network ensemble model and the third is boosting neural network ensemble model. The results indicate that boosting neural network ensemble model yields the best classification results. All three models show that projects with better preproject planning (lower PDRI score) are more likely to have a better chance of achieving project success. This paper provides some analytical results for the industry practitioners to put more efforts into the early planning in order to achieve better project performance.

References

[1] Breiman, L. (1996). "Bagging predictors." Machine Learning Journal, 24(2), 123-140

[2] Cho, C. S. (2000). "Development of the Project Definition Rating Index (PDRI) for Building Projects." Ph.D. Thesis, University of Texas at Austin, Austin, TX

[3] The Construction Industry Institute (1995). "Pre-Peoject Planning Handbook". Special Publication 39-2, Austin, TX

[4] The Construction Industry Institute (1996). "Project Definition Rating Index (PDRI) – Industrial Projects". Implementation Resource 113-2, Austin, TX

[5] The Construction Industry Institute (1999). "Project Definition Rating Index (PDRI) – Building Projects". Implementation Resource 155-2, Austin, TX

[6] Elhag, T. M. S., and Boussabaine, A. H. (2002). "Tender price estimation using artificial neural networks. II: Modelling.". J. Finan. Mange. Property Constr. 7(1), 49–64.

[7] Elhag, T. M. S. (2004). "Cost modeling: Neural networks vs. regression techniques." Int. Conf. on Construction Information Technology (INCITE), Langkawi, Malaysia, Construction Industry Development Board Malaysia (CIDB).

[8] Freud, Y. and Schapire R. E. (1996). "Experiments with a New Boosting Algorithm." Proc. International Conference on Machine Learning, San Francisco, 148-156

[9] Gibson, G. E. and Hamilton, M. R. (1994). "Analysis of Pre-Project Planning Effort and Success Variables for Capital Facility Projects", A report to the Construction Industry Institute, Source Document 105, Austin, TX

[10] Gibson, G. E. and Dumont, P.R. (1996). "Project Definition Rating Index (PDRI)" A report to the Construction Industry Institute, Research Report 113-11, Austin, TX
[11] Hansen, L. K. and Salamon P. (1990)." Neural Network Ensembles" IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(10), 993-1001
[12] Merrow, E. W., and Yarossi, M. E. (1994). "Managing Capital Projects: Where Have We been – Where are We Going?" Chemical Engineering, October: 108-111
[13] Roiger, R. J. and Geatz, M. W. (2003). "Data mining: a tutorial-based primer." Addison Wesley, Boston.

A Research Model for Architectural Meetings to Support the Implementation of New Building Technologies through Collaboration of Brainpower

Frans J.M. van Gassel, Ger J. Maas and Johanna E.M.H. van Bronswijk

Eindhoven University of Technology, The Netherlands, f.j.m.v.gassel@tue.nl

Abstract

Implementing new technologies in the domain of "Robotics and Automation in Construction" is necessary to enhance the values for clients and society. The technologies are available but implementing them is the coming challenge.

Building technologies are not used exhaustively or intelligently. In the manufacturing industry of complex products collaboration of brainpower through involvement of all stakeholders has resulted in a major enhancement of the values mentioned. Making the experiences of the involved professionals collaboratively explicit by well-prepared meetings can be a success factor to implement new building technologies. The aim of this paper is to develop a successful approach for organizing collaborative architectural meetings of the bonding type. This study has been divided into three phases: (i) getting insight in the relevant factors of a successful collaborative architectural meeting (desk research), (ii) analyzing case studies, and (iii) developing a meeting model for further research. The following parameters can describe an architectural meeting: (i) system variables, (ii) input variables, (iii) leading variables, (iv) participant variables, (v) tool variables, and (vi) outcome variables. These are included in a meeting model for further study to facilitate the collaboration of brainpower for enhancing the values for clients and society during the construction and use of buildings objects.

Keywords: building technologies, architectural meetings, collaborative brainpower

Introduction

Currently required values for clients of building projects include short building time and little waste, while values for society include safe working conditions and (social) sustainability. New building technologies to enhance these values for clients and society, such as automation and robotics, do exist, but are not implemented as quickly as desired (Hasegawa, 2006). Individual problem solvers, such as architects or building engineers, fall short in delivering these new values (Bennis and Biederman, 1997). To implement productively new building technologies in the architectural, engineering and construction industry specific management and social competences are needed and have to be learned. This could be reached by collaboration of brainpower. During architectural meetings collaborative involvement of all stakeholders is needed.

In other domains, such as the manufacturing industry of complex products (planes, computers, pharmaceutics), collaborative involvement of all stakeholders has resulted in a major enhancement of the values mentioned above (Sonnenberg, 2004). The ideas and concepts for these products were generated during creative meetings, involving all relevant professionals and even clients. Applying creative problem solving techniques (Osborn, 1963), energizing minds by engaging the senses (Roos, 2006) and team dynamics insights (Seagal et al, 2006) are proven instruments for successful implementation of innovative products.

Today's processes to create new products in the construction industry differ from those of the manufacturing industry by adherence to craft traditions based on time tested experience (Kvan, 2000). In addition, the various building professionals have their own language, symbols, unwritten rules, tools and paradigms (Buciarelli, 2002). Team forming is considered impossible because of a large turnover of labor. Professionals in building are rather solution driven and not problem driven, and show poor socio-emotional skills in their work (Zou et al., 2006; Glunk et al., 2008). These special aspects have to be considered by attempting to organize collaborative architectural meetings.

In 2006 Keursten et al. published a frame work for the building domain with elements that play a role in developing and using knowledge based on the "Corporate Curriculum" of Kessels (1996) that was highlighted in 16 case studies. But their approach did hardly realize the collaboration of brainpower during architectural face-to-face meetings.

Architectural meetings are a forum for reviewing functions, discussing unsolved issues, making decision, monitoring, and communicating decisions to those who need to know. Emmitt (2007) distinguishes five types of meetings: (i) controlling, (ii) coordinating, (iii) appraising, (iv) bonding, and (v) resolving.

In this paper we focus on the bonding type: "… meetings that fulfill a fundamental human need to communicate and bond, and hence foster team relationships. They create a sense of belonging and reflect the collective and cultural values of the temporary project organization" (Emmitt, 2007, p 17). The aim of this paper is to develop a successful approach for organizing collaborative architectural meetings of the bonding type.

Methodology

Our study has been divided into three phases: (i) getting insight in the relevant factors of a successful collaborative architectural meeting (desk research), (ii) analyzing case studies, and (ii) developing a research model for meetings.

The literature survey on collaborative brainpower as applied to the building process concerned (i) successful meeting types, (ii) their possible outcome, (iii) useful collaborative insights, (iv) reported success factors. In addition to searching literature databases (Webspirs, Web-of-Science, Scopus), the following journals were studied (1995-2007): Automation in Construction, Creative Problem Solving, Creativity and Innovation Management, Journal of Learning and Intellectual Capital, CoDesign, International Journal of CoCreation in Design and the Arts, Architectural Engineering and Design Management. Search key words used, included Creative Problem Solving (CPS), Collaborative design, Meetings, Creative thinking, Architectural management, and Meeting dynamics.

The collected information was structured according to Sebastian (2007) who developed a model for collaborating design-actors creating design-solutions. It involves three kinds of frames: (i) social (social environment, team work, behavior), (ii) cognitive (creativity, knowledge, decision), and (iii) project oriented (goal, vision, constraint, result).

In the second phase of the study a total of 15 meetings, organized by the first author, were studied by observation. In 12 of the 15 meetings analyzed the first author facilitated and the participants were mostly experts. The following parameters were assessed during observation: input, outcomes, control and mechanism of the activity meeting. The notation of the Systematic Analytical and Design Techniques was used (SADT, 2008). The aim of the meeting is considered the input variable. In 6 meetings the aim was "introduction in skills", in 7 meetings "vision building", and in 2 "vision implementation".

Two groups of interventions are distinguished: organizational interventions (1=not applied, 2=hardly applied, 3=strongly applied), and special interventions to explicate the experiences of the participants (0=not applied, 1=applied).

Rating of the outcomes of the meetings included providing, sharing and developing new knowledge (1=hardly any result, 2=mediocre result, 3=good result). However, the rate of developing new knowledge could not be analyzed because of too many missing data.

In the third phase the obtained data were structured and analyzed to support an iterative process leading to a collaborative architectural meeting research model.

Results

Relevant factors and insights

Successful bonding meetings appear to be a result of awareness, formulation and state of the needs of the (end) user and society (Kjølle 2005). Three types of outcome are distinguished after Bloom (2008) with the participants (i) knowing more and able to recall data or information, (ii) acquiring more skills, and (iii) developing the attitude to act as an expert. During architectural meetings cognitive processes take place: "To the thinking and interactive skills of design like the perceptive, the creative, the communicational, the learning and also the emotional and the teamwork processes" (Tschimmel, 2004). Relevant factors are numerous (Table 1), and form a source of inspirations cq interventions for the organizer to prepare an architectural face-tot-face meeting with a desired outcome.

Table 1. Insights in meeting processes, sorted according to the frame work of Sebastian (2007)

Frames	Relevant factors and insights
Social (social environment, team work, behavior)	- Socialization and externalization of implicit knowledge (Nonaka and Takeuchi, 1995) - To obtain desired interactions, the group had to work alternately in a generative and focusing mode (Hohn, 1996). - There are some studies about organizing architectural meetings. Gorse (2002) measured by the method of Bales the interpersonal communication and group interaction during construction management and design team meetings. Successful project outcomes were from groups who use a broader range of communication acts. "The level of positive emotion (agreeing and being supportive) is greater in successful teams".
Cognitive (creativity, knowledge, decision	- Learning styles and the circle of learning: concrete experience, observation and reflection, forming abstract concepts and testing situations (Kolb, 1983). These styles are incorporated in the tools variables of the meeting research model. See *Figure 1*. - Designing is learning (Dorst, 1997). - Unconscious thinking, rituals and creativity (Dijksterhuis, 2007). - Specific activity rhythm of divers personality dynamic during a face-to-face meeting (Seagal and Horne, 2006)
Project (goal, vision, constraint, result).	- The essence of collaborative design: collaboration from cooperation and coordination as a closely coupled design processes (Kvan, 2000). - Systematic Inventive Techniques (SIT) (Horowitz & Maimon, 1997). - The dialogue between intuitive thinking and the rational, logic thinking is the engine for the creative process (Groeneveld, 2008). - Reformulating the problem definition (Basadur, 2002). What is the real problem of the (end) user and what problem is interesting to solve? - A model with the following activities: formulating, representing, moving, bring problems and solutions together, evaluating and reflecting (Lawson, 2006). - The design process can divided in sub-processes as naming, framing, moving and reflecting (Valkenburg, 2000). - The development of a group to a team by passing trough the phases of forming, storming, norming, performing and adjourning (Robbins, 2002). - Van Gassel (2005) did experiments on the subject collaborative design by constructing metaphoric objects. The designers in this experiment were more involved but to a certain extent. To get the designers more involved in the design process tactical activities has to be added such story telling and recording on video to hold the ideas and make them communicable.

Case analysis

During the meetings a number of control activities and mechanism activities took place and were rated by the extent of its use (Table 2).

As to outcomes (rated 1 to 3), the mean for providing knowledge was 2.7, and for sharing knowledge 2.4. The lowest outcome (1=hardly any result) occurred in one out of the 15 cases analyzed.

To get some more insight in the cases two of them are described below.

Case "Awareness"

Participants: About 20 professionals as directors, designers, planning engineers, purchasers and project managers from the domains civil, road-building, housing and installing.

Description: To get more assignments a contractor needs a better dialogue with the client to create innovative bidding proposals. The management is not aware of the needed competences of brainpower development. A skills-learning meeting with the professionals was the aim to enhance more awareness, experience the required creativity skills and the creativity techniques that can be used for a certain outcome. Result: a number of professionals want to enhance their competences on brainpower development.

Assessment: this case obtained a rating of 3 for its outcome

Case "London Eye"

Participants: A group of 12 experts such as architects, project managers, cost accountants, facility managers and design managers.

Description: A Dutch contractor has been invited by the government to participate in a PPP bidding process for building a reception centre for asylum seekers at the airport of Schiphol. The contractor manager of the bidding activities wanted the participated professionals to focus on the client's wishes and dreams by using a metaphor. This metaphor has been developed during a day meeting. The meeting consisted of the system activities, and the tools brain writing, storytelling, selection techniques, making image boards of metaphors, reflection on the participant's contribution, etc. The activities were done in groups of two persons and plenary.

Table 2. Intervention scores during the 15 architectural meetings and the extent of their use. The organizational interventions were rated as: 1=Not applied, 2=Hardly applied, 3=Strongly applied. The special interventions were rated as: 0=Not applied, 1=Applied

Interventions		**n**	**Intervention scores**		
			Min	**Max**	**Mean**
Organizational activities	Socializing	14	1	3	1.7
	Acting	15	1	3	2.3
	Perceiving	15	1	3	2.1
	Incubation	15	1	3	1.5
	Reacting	15	1	3	1.8
	Receiving Feedback	14	1	3	1.8
	Imaging Problem	14	1	3	2.3
	Reformulating Problem	14	1	3	1.9
	Diverging	15	2	3	2.8
	Clustering	14	1	3	2.4
	Converging	14	1	3	2.1
Special activities to explicate the experience of the participants	Time out	15	0	1	0.1
	Splitting group	15	0	1	0.7
	Individual work	15	0	1	0.1
	Failures	15	0	1	0.1
	Rules	15	0	1	0.3
	Simulation	15	0	1	0.1
	Constructing	15	0	1	0.3
	Reflecting	15	0	1	0.1
	Serious Play	15	0	1	0.3
	Energizer	15	0	1	0.3

Result: the metaphor "London Eye" that describes the required values e.g. proud, identification, international, simplicity, daring en vision, and that is successfully used in the kick-of meeting.

Assessment: this case obtained a rating of 3 for its outcome

Towards a meeting model for further research

Tabel 2 shows us that organizational interventions were applied frequently and that these could be considered as system variables. In addition, Table 2 shows that special interventions were used incidentally and are considered by us as tool variables.

Based on both the literature survey and the case studies, the following variables appear to be most relevant for describing the activities of an architectural meeting.

(i) System variables (imaging the aim, reformulating the problem/question, diverging, clustering and converging);
(ii) Input variables (the aim of the meeting);
(iii) Leading variables (type of leader and time span);
(iv) Participants' variables (type of participants); and
(v) Tools variables (rational thinking, intuitive thinking, doing, dreaming, reflecting with an individual, small group and plenary appeal) , and
(vi) Outcomes variables (knowledge, skills and attitude).

Based on the results in Table 1 the variables are split up into sub-variables.

In Figure 1 these variables are depicted following the SADT notation and form together the meeting research model.

Discussion

The research model is restricted to meetings in the beginning phases of the building process under the architectural circumstances of the Netherlands, since the developed approach is based on a number of special preconditions of architectural meetings:

(i) Only meetings in the beginning of the building process are included; here there is more need to find the required client values than to find solutions.
(ii) Face-to-face meetings with widely divergent building experts.
(iii) Considerable turnover of participants in meetings. Long term team building is no issue.

In addition our approach has elements that deliver new knowledge about architectural meetings after validation:

(i) A well-founded organization of the meeting is an approach to get a better productivity of collaborative brainpower.
(ii) Defines minimal system activities for a meeting.
(iii) Uses well chosen tools to reach specific outcome.
(iv) Uses well chosen tools to anticipate on the specific contribution of the expert and ongoing meeting process.
(v) Is based on the process of experiential learning.

Keursten et al. (2006) found the following development principles for the support of learning processes leading to improvement and innovation, among other, enhance reciprocal appeal, search for passion, and temping towards knowledge productivity . Also some of his 10 design principles have a relation with our tool variables: formulating urgent and fascinating questions, and creative learning: knowledge developing by collaborative constructing.

The study resulted in an architectural meeting model of the bonding type with a set of variables describing input and outcome, as an useful structure for further observation and analysis to deliver syntax relations supporting the organizer of an architectural meeting to successfully set a meeting to deliver specific outcomes.

Based on our results we can formulate the following successful approach. An architectural meeting needs a minimal system configuration of the activities (i) imaging the aim, (ii) reformulating the question, (iii) diverging, (iv) clustering and (v) converging, and an alternating and a well-balanced variety of tools such as rational and intuitive thinking, doing, dreaming, reflecting and involvement of the whole group.

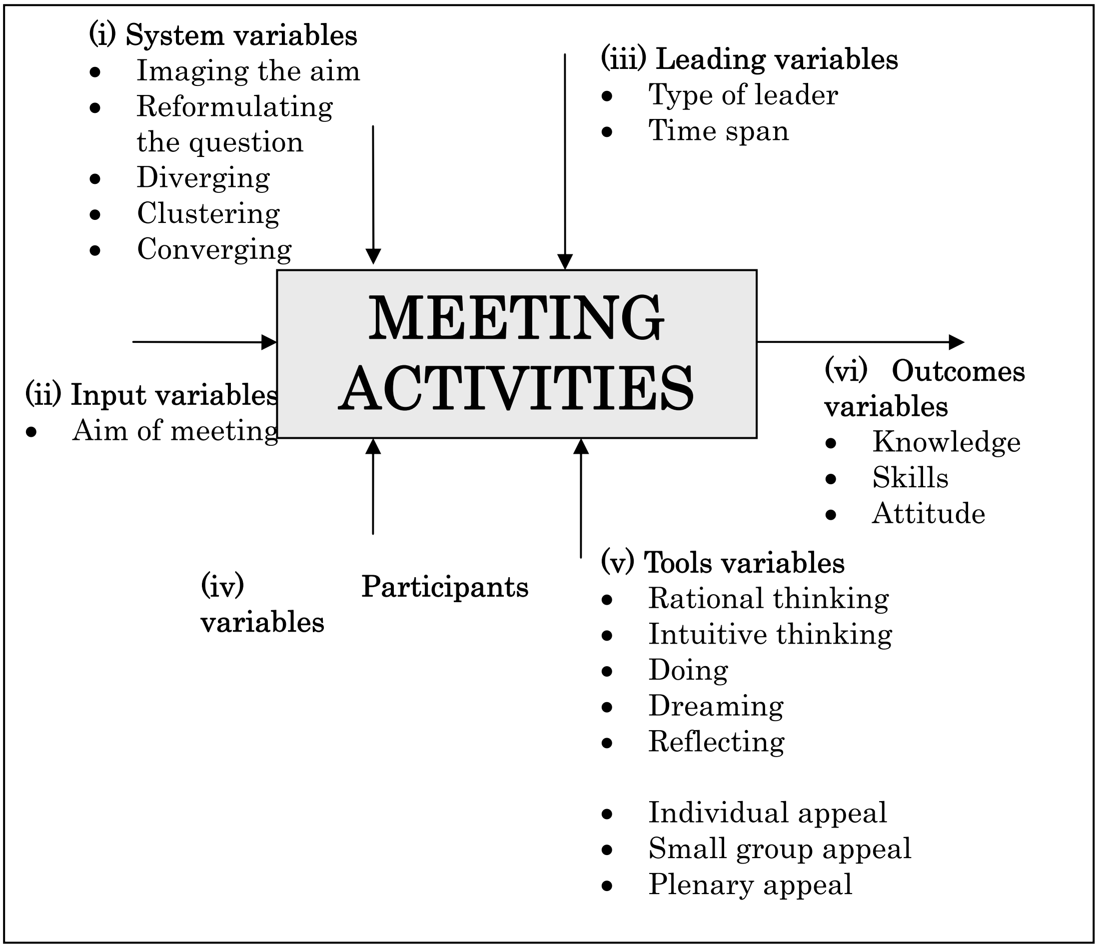


Figure 1. An architectural meeting model for further research as presented in SADT (2008) notation

References

[1] Basadur, M. (2002) "Challenge mapping for conceptual thinking – linking strategy to tactics". www.basadurSimplex.com Last consult 25th November 2008.
[2] Bloom B.S. http://en.wikipedia.org/wiki/Bloom's_Taxonomy Last consult 25th November 2008.
[3] Buciarelli, L. (2002) "Designing engineers". The MIT press.
[4] Dijksterhuis, A. (2007) "Het slimme onbewuste". Bert Bakker.
[5] Dorst, K. (1997) "Describing design: a comparison of paradigms". PhD thesis. Delft University of Technology, the Netherlands.
[6] Emmitt, Stephen (2007) "Design Management for Architects". Blackwell Publishing.
[7] Van Gassel, F. (2005) "Experiences with collaborative design by constructing metaphoric objects". In: Henri Achten (Ed) Design research in the Netherlands. Eindhoven University of Technology. Bouwstenen 92. 63-70.
[8] Glunk, U and R. Olie (2008) De sleutel tot succes (cultuur, samenwerking en innovatie in de bouw) PSIBouw.
[9] Groeneveld, R (2008) "De innerlijke kracht van de ontwerper". PhD thesis. Delft University of Technology.
[10] Gorse , C.A. (2002) "Effective interpersonal communication and group interaction during construction management and design team meetings". PhD thesis, University of Leicester.
[11] Hasegawa, Yukio (2006) "Construction automation and robotics in the 21st Century". In: N. Kano (Ed) Proceedings ISARC 2006 Tokyo.
[12] Hohn H. (1996) "Playing leadership and team development to innovative teams: a reflection on theory confronted with the perspective of experienced leaders". Delft University of Technology, the Netherlands.

[13] Horowitz, R and O. Maimon (1997) "Creative design methodology and the SIT method". In: proceedings of DETC'97.
[14] Kessels, J.W.M. (1996). "Knowledge productivity and the corporate curriculum". In J.F. Schreinemakers (Ed.). Knowledge management, Organisation, competence and methodology (pp 168-174) Wurzburg, Ergon Verlag.
[15] Keursten, P., S. Verdonschot, J. Kessels and K. Kwakman (2006). "Relating learning knowledge creation and innovation: case studies into knowledge productivity". International Journal of Learning and Intellectual Capital, 3 (41, 405-420).
[16] Kolb, D.A. (1983) "Experiential learning. Experience as the source of learning and development". Prentice Hall.
[17] Kvan, T (2000) "Collaborative design: what is it?". In: Automation in Construction 9 409-415.
[18] Nonaka, I. and H. Takeuchi (1995) "De kenniscreërende onderneming". Scriptum Management.
[19] Osborn, A.F. (1963) "Applied imagination: principles and procedures of creative problem-solving". New York, Scribner's Sons.
[20] Robbins, S.P. (2002) "Gedrag in organisaties". Prentice Hall.
[21] Roos, J. (2006) "Thinking from within. A Hands-on Strategy Practice", Palgrave Macmillan, New York.
[22] SADT http://en.wikipedia.org/wiki/Structured_Analysis_and_Design_Technique Last consult 25th November 2008.
[23] Seagal, S and D. Horne (2006) "Human Dynamics". Scriptum Management.
[24] Sebastian, Rizal (2007) "Managing Collaborative Design". PhD thesis. Delft University of Technology, Eubron.
[25] Sonnenburg, S. (2004) "Creativity in Communication: A Theoretical Framework for Collaborative Product Creation". Creativity and Innovation Management, Volume 13, number 4.
[26] Tschimmel, Katja (2004) "A new discipline in design education: cognitive processes in design". In: Proceedings International Engineering and Product Design Education Conference, Delft.
[27] Valkenburg, R. C. (2000) "The reflective practice in product design teams". PhD thesis. Delft University of Technology, the Netherlands.
[28] Zou, P.X.W. and H. Darvish (2006) Group Assignments and Teamwork Skills Development in Postgraduate Construction Mangement Studies. In: Architectural Engineering and Design Management 2 (2006) 203-215.

A social capital perspective to innovation management in construction

Remy D. van der Vlies[1] and Ger J. Maas[2]

1PhD candidate at Eindhoven University of Technology (TU/e), Faculty of Architecture, Building and Planning, Performance Engineering for Built Environments (PEBE) and consultant at Centrum voor Innovatie van de Bouwkolom, Zeist, The Netherlands, r.d.v.d.vlies@tue.nl
2Prof. Performance Engineering for Built Environments (PEBE) at Eindhoven University of Technology (TU/e), Faculty of Architecture, Building and Planning and Director of Strategy of Royal BAM Group.
P.O. Box 513, 5600 MB Eindhoven, the Netherlands, g.j.maas@tue.nl

Abstract

State-of-the-art products commonly outperform construction products that are used in day to day building practice. Also construction products appear to have a diffusion curve that differs from consumer products, slower at first, more rapid later. The social capital theory helps us to understand why certain actors are able to get their ideas adopted and why others do not. Aim of this paper is to explore to what extend social capital theory may provide explanations for the way in which innovative construction products are diffused. Therefore social capital literature and building process literature are compared.

Construction process literature shows that the industry is fragmented and contacts between the various professional networks are limited to that in the building projects. Even more so the contact in construction projects appears to be short-term oriented.

Social capital literature provides an explanation (network closure) for both opportunistic behavior in construction projects as well as lack of reward for those who put in extra effort when adopting an innovation. Also social capital literature shows that those who are able to bridge the gaps between networks (structural holes) are able to get their ideas adopted more easily and so are able to spread their innovations more rapidly. Social capital theory thus appears to be helpful to explain how the diffusion of building product innovation can be improved.

Introduction

Innovative products commonly outperform construction products that are used in day to day building practice. Market introduction of innovative products enhances the competitive position of a firm (Seaden and Manseau 2001). Innovative construction products appear to have a S-shaped diffusion curve that differs from consumer products, slower at first, more rapid later (Lichtenberg 2002).

S-curves are a model of reality, made to help us understand a reality that is otherwise too complex to understand. The S-curve is a plausible way to parameterize the diffusion process and is roughly consistent with the facts. 'Roughly consistent' because diffusion curves tend to be asymmetric in practice and also because of the fact that most innovations fail (i.e., they do not diffuse at all). (Geroski 2000)

What are the prime influences on the shape of the diffusion curve? Six primary influences driving or hindering construction innovation were identified by Blayse and Manley (Blayse and Manley 2004):

1. clients and manufacturers;
2. procurement systems;
3. regulations/standards;
4. the structure of production;
5. relationships between individuals and firms within the industry and between the industry and external parties;
6. the nature and quality of organizational resources.

However, the relationship between these influences themselves and with other aspects of business strategy is unknown. (Blayse and Manley 2004). From the perspective of a single business in the construction industry only the last three can be influenced. This research therefore further investigates the structure of production, relationships between individuals and firms within the industry, and the nature and quality of organizational resources enhancing the diffusion of construction innovation.

For this research, the definitions for innovation used by Rogers (Rogers 1995) and Emmitt and Yeomans (Emmitt and Yeomans 2008) can be combined and re-written as: *An innovation is a building product that is perceived as new by a principal, specifier, or engineer.* The diffusion of an innovation is perceived as: *the process by which innovation is communicated through certain channels over time among the members of a social system* (Rogers 1995).

Social capital theory helps us to understand why an actor is able to get his ideas adopted and why some other can not (Burt 2004), and thus social capital may help us to understand why some construction firms are able to get their innovations diffused and why others do not.

All social relations and social structures facilitate some forms of social capital (Coleman 1988). Like physical and human capital, social capital is a productive resource. Social capital facilitates a firm's business operations. (Tsai & Ghoshal 1998). Nahapiet and Ghoshal (Nahapiet and Ghoshal 2005) define: *Social capital is the sum of the actual and potential resources embedded within, available through, and derived from the network of relationships possessed by an individual or social unit.*

It appears to be a useful perspective to help manage the diffusion of innovation.

The aim of this paper is to explore to what extent social capital theory may provide explanations for the way in which innovative construction products are diffused. To do so, literature on social capital theory and building process characteristics are compared.

Method

Literature is searched using both Google Scholar and the literature database ISI Web of Knowledge. Key words used for identifying construction process characteristics are:

- “characteristics construction process”
- “construction supply chain integration”
- “life cycle contracts construction”
- “technology diffusion construction innovation”

Key words used for searching descriptions of social capital theory.

- “definition social capital”
- “definition structural holes”
- “definition network closure Coleman”
- “slow diffusion; social network”
- “brokerage diffusion”

Results

Construction process characteristics

Construction process literature shows that (i) the industry is fragmented and (ii) contacts between the various professional networks are limited to that in the building projects. Furthermore, (iii) the contact in construction projects appears to be short-term oriented.

(i) Fragmented

There are many relatively small companies in the construction industry. In the Netherlands, almost 90% of all building companies have no more than ten employees, and almost 10% of the companies are medium-sized firms with from ten to 100 employees (Priemus 2004). Dulaimi e.a (Dulaimi et al. 2002) conclude that the fragmentation and especially the segregation of design and construction activities are the main barriers to an improved performance of the industry.

(ii) Project-based contacts

Dubois and Gadde (Dubois and Gadde 2002) studied the operations and behaviors of firms as a means to dealing with complexity. Industry as a whole is featured as a loosely coupled system. The pattern of couplings builds on two interdependent layers: tight couplings in individual projects and loose couplings based on collective adaptations in the permanent network. They concluded that the characteristics of the industry seem to favor short term productivity while hampering innovation and learning. Van Hal (van Hal

2000) concluded that one of the prime factors hampering the diffusion of innovation in the housing industry is the lack of information transfer between projects. This information transfer must be based on unambiguous evaluations and an innovation champion is needed in order to implement the innovation successfully.

(iii) Short term orientation

Vrijhoef en Koskela (2000) characterize the construction process as short term oriented. In contrast to manufacturing systems, where multiple products pass through the factory, and are distributed to many customers, the `construction factory' is set up around the single product. Therefore every project creates a new product or prototype. In general there is little repetition. As a result, the construction supply chain is typified by short term organizations, instability, fragmentation, and especially by the separation between the design and the construction of the built object. Kumaraswamy (KUMARASWAMY 1998), wonders why the construction industry's short-term orientation did not lead to innovative managerial techniques.

Key innovations actors

Seaden and Manseau (2001) defined no less than ten key actor types involved in construction who can undertake innovation activities:

1. Building materials producers,
2. Machinery manufacturers,
3. Building product component manufacturers,
4. Sub-assemblers (trade specialty and installers),
5. Developers and facility assemblers (or main contractors) ,
6. Facility/building operators who manage property services and maintenance,
7. Architects and specifiers,
8. Consultants and engineers,
9. Providers of complementary goods and services such as transportation, distribution, cleaning, demolition and disposal.
10. Institutional environment actors such as financial institutions and business/trade general labour regulations and standards.

Vrijhoef and Koskela (Vrijhoef and Koskela 2000) drew a supply chain based on a traditional construction supply chain, see figure 1. Based upon this supply chain one may identify two more key actors:

1. Residents,
2. Principals.

The total of different types of actors in the construction industry added up twelve.

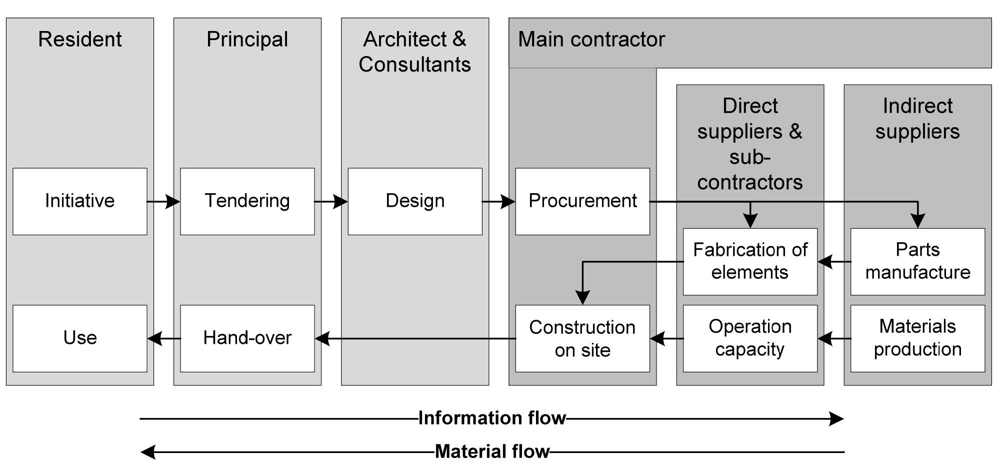


Figure 1: Typical configuration of the traditional construction supply chain (Vrijhoef and Koskela 2000)

Integrated contracts

The characteristics mentioned so far specifically describe the traditional construction process. Since the 1990's however, several alternatives to the traditional construction process came to market (Briscoe and Dainty 2005). Amongst those are the design and build contract, the private finance initiative (PFI), the public private partnership (PPP), and the build, own, operate and transfer (BOOT) (Briscoe and Dainty 2005; Brady, Davies and Gann 2005; (Ndekugri and Turner 1994)).
Design and Build contracts reduce fragmentation. Figure 2 shows the typical configuration of its supply chain.

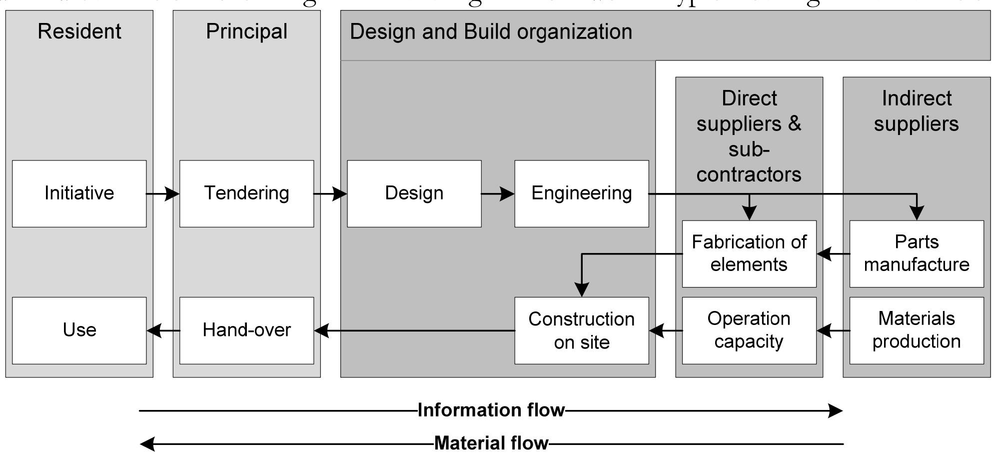


Figure 2: Typical configuration of the design-and-build supply chain (based on: Vrijhoef and Koskela 2000)

The growth in projects procured under the PFI, the PPP, and under BOOT arrangements should result in less short term orientation. Where even maintenance services are part of the contract, the focus will probably be on the complete lifespan of the building. (Briscoe and Dainty 2005)

(Brady, Davies, and Gann 2005) found that the construction industry perceptions of value, systems integration and integrated solutions suggest that the integrated and long term contracts are still at an early stage in its development. They suggest that best opportunity for the introduction of integrated and long term contracts is in the context of private finance initiatives in the public sector or large clients who require repeatable solutions in the private sector.

Social capital theory

Social capital theory helps us to understand why certain actors are able to get their ideas adopted and why others do not (Burt 2004). The term social capital was originally used to describe the relational resources, embedded in cross-cutting personal ties. Later research has applied the concept to a broader range of social phenomena, including relations within and beyond the firm. (Tsai and Ghoshal 1998). For long there have been two approaches to social capital: (i) The network closure approach focuses on the density of the network. Whereas (ii) the structural holes approach focuses on the gaps between several dense networks.

(i) Network closure approach

Coleman (Coleman 1988) considered network closure to be conditional for the existence of obligations, expectations, and social norms within a network.

These arise as attempts to limit negative actions of actors or encourage positive ones. In an open network like that of figure 3a, actor A, having relations with actors B and C, can carry out actions negative for B or C or both. Since they have no relations with each other, but with others instead (D and E), they cannot combine forces to sanction A in order for him to stop his actions. Unless either B or C alone is sufficiently harmed and sufficiently powerful to sanction A alone. In a network with closure, like that of figure 3b, B and C can combine to provide a collective sanction, or either can reward the other for sanctioning A.

(ii) Structural holes approach

Structural holes may exist as either empty spaces or as negative ties between alters and/or groups. Both types of structural holes act as buffers, like an insulator in an electric circuit… people on either side circulate in different flows of information (definition by (Oliver, Kalish, and Yair 2007) based on a definition of Burt).

Those that are able to circumnavigate the structural holes are the ones that are able to combine ideas from both networks into an invention or enable innovations to flow from one network to the other. According to Burt (2004) it is the intermediate actor, or broker, that provides social capital:
Opinion and behavior are more homogeneous within than between networks, so people connected across networks are more familiar with alternative ways of thinking and behaving. Brokerage across the structural holes between networks provides a vision of options otherwise unseen, which is the mechanism by which brokerage becomes social capital (…) The between-networks brokers are more likely to express ideas, less likely to have ideas dismissed, and more likely to have ideas evaluated as valuable.

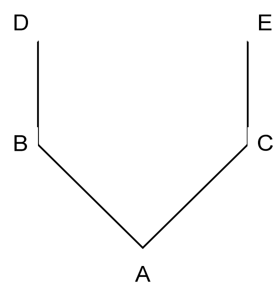


Figure 3a: Network without closure

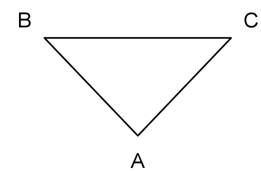


Figure 3b: Network with closure (Coleman 1988)

Diffusion of innovation

Both network closure and structural hole theory view reciprocity as the mechanism that turns relationships into the assets that define social capital (Gargiulo and Benassi 2000). Burt (Burt 2000) integrated them into one integrated theory: *Structural holes are the source of value added, but network closure can be essential to realizing the value buried in the holes. (p…)*

Network closure may sometimes hamper the diffusion of innovation. For instance:
*The more a manager was strongly tied to a cohesive group of peers, the less able he or she was to adapt his or her communication network to the changes brought about by the global organizational change (…).*Gargiulo & Benassi (1998, p…)

On the other hand, network closure may also enable the diffusion of an innovation. Rogers (1995) identified, compatibility of an innovation with the norms and values of an actor as an important factor enabling their adoption of the innovation. Therefore, the diffusion of innovation is easier amongst a network with high closure.

So, depending on the values and norms within a network, closure can either hinder or enhance the diffusion of an innovation.

Spencer (Spencer 2003) showed an association between a firm's status as a global knowledge broker in the last period before dominant design emergence and its ability to successfully make the transition into commercial production by installing large-scale manufacturing facilities. Brokers apply, filter, and reframe knowledge as they pass it on. And this intentional or unintentional reframing may help innovating firms shape the emerging institutional environment to favor the diffusion of their own technologies.

Conclusion: fragmentation vs. integration, short-term orientation vs. life cycle approach, structural holes and network closure in the construction industry.

From a social capital perspective the construction industry is a large sector with little network closure and many structural holes. The lack of network closure results in the short term orientation. The fragmentation and project based contacts result in many structural holes.

The lack of network closure enables actors in the construction industry to aim for short term benefit and behave opportunistic. Because of the combination of 1) temporary organizations and 2) fragmentation, the chances that actors B and C (see figure 3a) are able to enforce their norms upon actor A are next to zero. Something that is even more difficult when it comes to principals. Many of whom build only once in their lifetime. The lack of network closure is not only the reason why opportunistic behavior is not punished. It also causes that putting in extra effort (which is most likely needed when adopting an innovation) is not rewarded.

At least twelve different key innovation actors may be identified which encounter each other only in temporary projects. However, there are some lasting networks in the construction industry, but these are amongst groups of peers. Architects meet each other in their architects association, as do engineers in their engineers association, builders in their builders association, etc, etc. The only interdisciplinary contact, or brokerage opportunity, exists within the construction project. Deroians' (Deroian 2002) statement that some innovations need delay to diffuse, others often fail seems to apply to product innovation in construction. He suggests that the formation of social networks explain the diffusion of innovation. Interaction is conceived as influence effects and the network of interpersonal influences is learning step-by-step. The gradual formation of the social network leads, after a period of latency, to a collective evaluation of the innovation. Given the large amount of structural holes this might explain why innovative construction products appear to have a diffusion-curve that differs from consumer products, slower at first, more rapid later. In the construction industry it takes longer for the social network to form around the many structural holes and come to the collective evaluation.

Social capital theory thus appears to be helpful to explain how the diffusion of building product innovation can be improved. In order to be able to actually manage the social capital of actors in construction further research is needed on both:

- Network closure: how can good conditions for the adoption of an innovation be created? What actors (fragments) appear to have the least short term orientation?
- Structural holes: how does the collective evaluation take place? How fragmented is the construction industry really?

References

[1] Blayse, A. M. and K. Manley, "Key influences on construction innovation," Innovation 4 (3): 1-12 (2004).
[2] Brady, T., A. Davies, and D. Gann, "Can integrated solutions business models work in construction?," Building Research & Information 33 (6): 571-579 (2005).
[3] Briscoe, G. and A. Dainty, "Construction supply chain integration: an elusive goal?," Supply Chain Management-An International Journal 10 (3-4): 319-326 (2005).
[4] Burt, R. S. 2000. The network structure of social capital. Vol. 22, 2000. NEW YORK: JAI-ELSEVIER SCIENCE INC.
[5] Burt, R. S., "Structural holes and good ideas," American Journal of Sociology 110 (2): 349-399 (2004).
[6] Coleman, James S., "Social Capital in the Creation of Human Capital," The American Journal of Sociology 94: S95-S120 (1988).
[7] Deroian, Frederic, "Formation of social networks and diffusion of innovations," Research Policy 31 (5): 835-846 (2002).
[8] Dubois, A. and L. E. Gadde, "The construction industry as a loosely coupled system: implications for productivity and innovation," Construction Management and Economics 20 (7): 621-631 (2002).
[9] Dulaimi, M. F. et al., "Enhancing integration and innovation in construction," Building Research and Information 30 (4): 237-247 (2002).
[10] Emmitt, S. and D. T. Yeomans. 2008. SPECIFYING BUILDINGS: A Design Management Perspective.Butterworth-Heinemann.
[11] Gargiulo, M. and M. Benassi, "Trapped in Your Own Net? Network Cohesion, Structural Holes, and the Adaptation of Social Capital," Organization Science 11 (2): 183-196 (2000).
[12] Geroski, P. A., "Models of technology diffusion," Research Policy 29 (4-5): 603-625 (2000).
[13] KUMARASWAMY, M. M., "Industry development through creative project packaging and integrated management," ENGINEERING CONSTRUCTION AND ARCHITECTURAL MANAGEMENT 5 (3): 228-237 (1998).
[14] Lichtenberg, J. 2002. Ontwikkelen van projectongebonden bouwproducten.Ph D thesis). Delft, 2002.
[15] Nahapiet, J. and S. Ghoshal, "Social capital, intellectual capital, and the organizational advantage," Sumantra Ghoshal On Management: A Force For Good (2005).
[16] Ndekugri, I. and A. Turner, "Building Procurement by Design and Build Approach," Journal of Construction Engineering and Management 120: 243 (1994).

[17] Oliver, Amalya L., Yuval Kalish, and Gad Yair, "Reflections on "Brokerage and Closure"," Social Networks 29 (2): 330-339 (2007).

[18] Priemus, H., "Dutch contracting fraud and governance issues," Building Research and Information 32 (4): 306-312 (2004).

[19] Rogers, E. M. 1995. Diffusion of Innovations.Free Press.

[20] Seaden, G. and A. Manseau, "Public policy and construction innovation," Building Research & Information 29 (3): 182-196 (2001).

[21] Spencer, J. W., "Global Gatekeeping, Representation, and Network Structure: A Longitudinal Analysis of Regional and Global Knowledge-Diffusion Networks," Journal of international bussiness studies 34 (5): 428-442 (2003).

[22] Tsai, W. P. and S. Ghoshal, "Social capital and value creation: The role of intrafirm networks," Academy of Management Journal 41 (4): 464-476 (1998).

[23] van Hal, A. 2000. Beyond the Demonstration Project: The Diffusion of Environmental Innovations in Housing.Uitgeverij Aeneas BV.

[24] Vrijhoef, R. and L. Koskela, "The four roles of supply chain management in construction," European Journal of Purchasing and Supply Management 6 (3-4): 169-178 (2000).

A Framework of Critical Resource Chain in Project Scheduling

S. S. Liu,[1] and K. C. Shih[2]

[1]Assistant Professor, Department of Construction Engineering, National Yunlin University of Science and Technology, No. 123, Section 3, University Road, Touliu, Taiwan, 640, ROC, Tel:886-5-5342601 ext. 4724, Fax: 886-5-5312049, Email:liuss@yuntech.edu.tw.
[2]Ph.D. Student, Graduate School of Engineering Science and Technology, National Yunlin University of Science and Technology, No. 123, Section 3, University Road, Touliu, Taiwan, 640, ROC, Tel:886-5-5342601 ext. 4701, Fax: 886-5-5312049, Email:g9310816@yuntech.edu.tw

Abstract

To analyze a schedule is beneficial to understand and organize a project. Project schedules, in which time plans are created based on the Critical Path Method (CPM) or on Resource-Constrained Project Scheduling Problem (RCPSP) optimization, are targets herein. According to the concepts of the Theory of Constraints (TOC), a schedule is treated as a system. Schedule elements are suspected constraints. Moreover, a goal depends on the policy of schedule creation. Schedules are further analyzed herein to recognize its bottlenecks. Additionally, resource information is surveyed to identify true constraints. To integrate identified constraints on a schedule, a framework is proposed and the concept named the critical resource chain is introduced. Three scenarios are employed to illustrate the proposed framework under different scheduling considerations. Issues of project understanding can then be explained and discussed through identified constraints.

1. Introduction

A schedule guides a project, and it is also an index of project performance. In order to create schedules, the Critical Path Method (CPM), which is a scheduling method familiar to most project planners, is employed. According to the concepts of CPM, schedules are created by arranging activities and identifying at least one Critical Path (CP). As an extension of CPM, the Resource-Constrained Project Scheduling Problem (RCPSP) which states that schedules are created under resource limitations has been investigated for years. Whether using CPM or RCPSP, optimization operations have been applied to solve scheduling problems. Thus, optimized schedules are possible to lead projects.

Although a schedule offers a sketch of a project with activities, activity information can not totally satisfy the requirements for project understanding. Further investigation on schedules is required. Thus, this study examines resource information, which is schedule elements other than activities, using the Theory of Constraints (TOC). Resource related constraints can be recognized, and resource importance can then be explained. Moreover, Issues of project understanding can be further discussed.

2. Theories

Necessary concepts of TOC and RCPSP optimization are reintroduced to support and lead to the proposed method.

2.1 TOC and critical chain

The theory of constraints (TOC), which was proposed by (Goldratt, 1984), is a problem analysis tool. Based on the concepts of TOC, problems can be analyzed in terms of three basic fundamentals: system, goal, and constraint(s). In order to refine a system, a goal must be set, and constraint(s) must be found. The entire analysis process of TOC can be defined as follow: 1. Identify constraint(s); 2. Decide how to exploit the constraint(s); 3. Subordinate everything else to the above decision; 4. Elevate the constraint(s); 5. If in any of the previous steps a constraint is broken, go back to step 1. Briefly, TOC analysis is a process of system refining by removing constraint(s) based on a set goal. Moreover, constraint identification is the key foundation to support the five-step analysis process.

CPM based schedules can certainly be analyzed by the concepts of TOC. According to CPM rules, schedules are built from activities, activity durations, and relationships between activities. These elements are considered as bases to create a schedule "system," and some of them may be "constraint(s)" withal. The "goal" of such a schedule system is minimized overall schedule duration because of a CP in which a group of activities are connected to organize an "irreducible" total activity duration. Constraints of schedules are clearly implied in CP activities. In order to apply TOC in project management, critical chain (CC) and buffer management are proposed.

CC is the CP where schedules are created with consideration of resource insufficiency. For example, if a mutual resource is required by several activities at the same time, some of these activities may be postponed till the resource is available again. Such a resource insufficiency results in adding finish-to-start activity relationships among activities. These activity relationships are named resource dependencies in TOC. Furthermore, the CP is renamed CC when at least one resource dependency occurs in the CP. In other words, resource dependency is considered a constraint if the term CC is adopted.

Buffer management is applied to risk issues. An activity buffer is identified as an extension of activity duration considering overall risk evaluation of an activity. As TOC proposed, activity buffers that belong to identical activity paths are integrated in a path buffer, and this path buffer is located at the end of a path. Furthermore, project buffer and feeding buffer are two types of path buffer that are introduced in TOC. A project buffer is the buffer of CP (CC), and feeding buffers which belong to non-CP (CC) paths. Moreover, buffers also indicate the degree of risk being considered in a schedule. As an example of project buffer, risks are ignored in a project if no project buffer exists. A long project buffer implies that risks are possible and there is concern they may cause delays in the overall schedule duration.

Schedules considering TOC should be created using exact activity durations, and risk issues ought to be integrated in buffers. However, several questions are not answered in TOC: how can planners create schedules given resource insufficiency and how to validate a resource dependency are not specified.

2.2 RCPSP optimization

RCPSP optimization offers solutions to the above questions. RCPSP optimization has been widely discussed and identified for project scheduling issues. (Herroelen, 1998) and (Brucker, 1999) collected and classified RCPSPs and proposed optimization models. RCPSPs inherit most CPM rules but not resource assumptions. First, resource limitations are concerned in RCPSPs. Activities may be postponed when a resource insufficiency occurs. Second, various productivity options are usually allowed for each activity. Each productivity option corresponds with a combination of resource usage to present related activity duration. A productivity option for each activity must be determined to create a schedule. Finally, RCPSP allows restrictions other than schedule elements if they are required. Thus, RCPSPs are more complex than CPM scheduling problems. In order to solve RCPSPs, optimization operations have been well investigated. Optimization operations are not further explicated herein since the reader can refer to references mentioned before.

Comparing RCPSP optimization with CPM scheduling, resource importance is concerned. Resources are more effective than activities for producing schedules. Though activities are still elements of schedules, activities cannot dominate an entire schedule as they do in CPM, such that activities are assumed to drive resources. Not only does resource availability affect start time of activities, but productivity options determine for how long an activity will be performed. Therefore, any unit of resource may influence an entire schedule.

Schedules based on RCPSP optimization can be foundations for TOC analysis. RCPSP optimization and TOC have many characteristics in common. When RCPSP optimization creates a schedule with an objective function, an equal goal exists in TOC. Furthermore, RCPSP optimization is also a schedule improvement process to approach perfect scheduling. Dominated schedules are screened out during optimization. Once an optimal solution is found, the refined schedule is assured. Thus, identification of constraints can be performed based on the secure five-step of TOC schedule.

Buffer management is adopted for all schedules in this study for risk issues. First, schedule elements can be estimated accurately in RCPSP optimization and CPM scheduling for exact schedules. Constraints can then be defined accordingly. Next, the length of a buffer can represent the overall evaluation of both risk possibility and effect as duration extension. Furthermore, project buffer and feeding buffers further integrate individual activity buffers. A project buffer and feeding buffers can offer an overview of the risk issues facing a project. In order to clearly define feeding buffer and project buffer in this study, float time of non CP (CC) as defined as in CPM and is integrated in feeding buffers. Moreover, a project buffer is equal to the project duration minus the overall schedule duration if contractual project duration exists.

To refine a schedule through the five-step TOC analysis, what are constraints is still the key question. Based on the concepts of CPM, CP (CC) activities imply constraints on a schedule. According to RCPSP optimization, activities and resources interact. Thus, resource information is required to recognize constraints.

3. Method

3.1 Constraint Identification

In order to recognize constraints for an optimized schedule, the minimized project duration is assumed as the goal used to identify constraints. Optimization is employed to schedule creation of both a CPM schedule problem and RCPSP to ensure secure schedules. Based on these assumptions, constraint identification can be discussed as follows:

1. Resource Limitation and Critical Resource

Resource limitation is a clue to constraints. Resource limitation can cause resource insufficiency and further constraints of resource dependency. However, resource dependency constraints have resulted in translation from CP to CC, and they cannot further contribute to TOC analysis. Thus, the reason for resource insufficiency, resource limitation, needs to be surveyed. By drawing resource diagrams of a created schedule, it can be investigated whether resource limitations can be labeled constraints whenever resource usage reaches resource limitation. Especially, these resources are employed by CC activities at that moment. If additional resources are considered, the schedule may have a chance to be refined. Therefore, resource limitation at that moment is considered as a suspected constraint. Moreover, Max usage of resources plays a role similar to resource limitation. Although CPM schedules assume unlimited resources, an oversupplied resource does not worth anything to a project, and the max usage of a resource has a high probability of being the final resource quantity of a project when performing schedules. Thus, Resource limitation and max usage are two criteria to identify suspected constraints herein.

Whenever a suspected constraint of resource limitation is recognized, this resource is identified as a critical resource. In order to differentiate suspected constraints, information on a critical resource contains when the constraint occurs, what is the corresponding resource, and information on resource limitation or max usage. Furthermore, although CP (CC) activities can extend through the entire schedule, not all resources are employed by CP (CC) activities. In other words, some critical resources are not constraints. Thus, critical resources need to be further defined as follows: Whenever a resource is totally occupied by CP (CC) activities, the term of superior critical resource is adopted. On the contrary, the term inferior critical resource is adopted whenever a resource is used up regardless of CP (CC) activities. The definitions of all types of critical resource are shown as Figure 1. Based on these definitions, superior critical resources must be constraints.

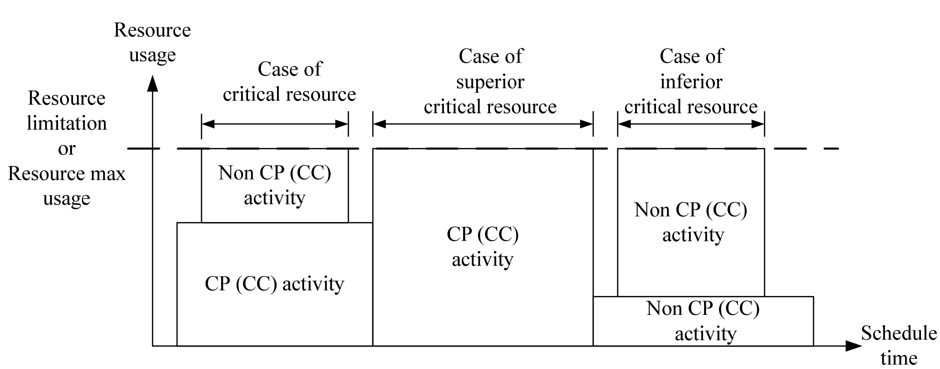


Figure 1. Sample of critical resource identification

2. Productivity Constraints

CP (CC) activities imply constraints. Although CP (CC) activities lead to minimized overall schedule duration, projects are not restrained by activity durations of CP (CC) activities but by productivity options. If a productivity option with a short corresponding activity duration is offered for a CP (CC) activity, the overall schedule duration can be directly shortened. Therefore, productivity options that CP (CC) activities adopt are treated as constraints herein. In order to present a productivity option constraint, activity name, corresponding activity duration, and quantity of resource usage are elements to differentiate other options. When an activity is identified as CP (CC) activity, the employed productivity option of the activity must be a constraint.

Productivity options imply additional constraints. It is a foundational assumption of CPM that resources offer a certain productivity for activities. That is why activity durations can be estimated according to resource usage. In other words, whether resources constitute a productivity option or even any individual resource is assumed to perform with steady productivity, they are natural constraints of schedules.

3.2 Critical Resource Chain Framework

A framework is proposed to integrate constraint information as figure 2. The framework is structured with all kinds of critical resources, productivity options for each CP (CC) activity, and project buffers. The concept of critical resource chain is then proposed since such resource related constraint information is shaped as a chain to restrain the entire schedule.

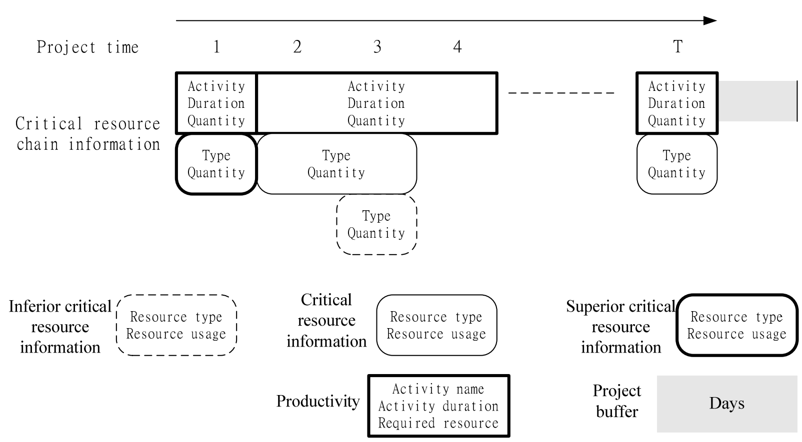


Figure 2. Sample of critical resource chain

3.3 Resource importance

Based on constraint identification as above, resource importance can be further discussed. First of all, a superior critical resource occurs with CP (CC) activities to indicate that a resource limitation directly restrains the schedule. The resource limitation is definitely a constraint on the schedule at that moment. In order to refine the schedule by removing this constraint, additional resources can be helpful. Contrarily, if any shortage occurs such as machine failure, the schedule will be also directly influenced. Thus, the resource that a superior critical resource indicates is a key resource. The required quantity of resource must be satisfied at that moment according to its classification as a superior critical resource. Second, critical resource status indicates the lower level of resource importance than superior critical status. In this case, there are chances to avoid impacts of resource variance by postponing non CP (CC) activities. Finally, inferior critical resource information is regardless of the goal. However, critical resources and inferior critical resources relate to other project management issues.

Resource productivity can also represent resource importance. In order to refine the schedule, considering various productivity options is beneficial. Although a selection of productivity options has been made during optimization, productivity options can be further examined in TOC analysis if a constraint on a productivity option is considered to be broken to refine the optimized schedule. When all resource productivity options of CP (CC) activities are scheduled, resources are crucial to offer their expected productivity with corresponding resource usage. Not only is resource usage of these productivity options required, but variance of productivity offering, especially a low productivity, must be prevented.

4. Scenario experiment

A case experiment is employed to implement the proposed critical resource chain framework. Two scenarios are considered as follows: Scenario 1 is a CPM scheduling problem. The minimized overall schedule duration is set as the goal, and resource limitation is not involved. Moreover, each activity adopts the shortest duration of productivity options to meet the goal. Scenario 2 represents a RCPSP optimization case with the goal as in scenario 1. Additionally, resource limitation is considered, and various productivity duration options are allowed. Comparing scenario 1 with scenario 2, constraint difference between CPM and RCPSP based schedules can be demonstrated. Project information and optimized output of these three scenarios are listed as Table 1.

Resource dependency behaviors can be easily discovered. The activity path A-C-G-H-I is the CP in scenario 1. The order of activities matches activity relationships to successors. The activity relationship A-C demonstrates that activity C is one of the successors of activity A as well as other relationships in CP are. In scenario 2, two activity paths, A-B-C-E-G-H-I and A-D-G-H-I, compose the minimized overall schedule duration: 41 days, and both these two paths can be CP or CC. Several resource dependency cases can be observed. B-C, C-E, E-G and D-G are additional activity relationships that are not set in project information. Thus, the CC is ascertained.

Various productivity options are beneficial and necessary in an RCPSP based schedule. By setting the goal of the minimized overall schedule duration, activities have no reason to adopt productivity options other than the one with the shortest duration. However, activities C, D, F, G, and H in scenario 2 adopt a productivity option other than the shortest one under the same goal. This situation is based on interactions of schedule elements. Nevertheless, the shortest activity duration of productivity options of each CP (CC) activity cannot guarantee the minimized overall schedule duration for sure in the RCPSP based schedule. Another investigation of activity H is significant. Activity H has two productivity options: 2-day case requiring $4,700 and 3-day case requiring $5,250. The 3-day case is dominated by the 2-day case no matter whether activity duration or usage cost is the measure considered. However, it is beneficial in scenario 2 to consider activity H as a CC activity. Restate, various productivity options is required while scheduling.

Critical resource chain frameworks of scenario 1 and 2 are implemented with related resource diagrams as Figure 3. Information of non CP (CC) activities is added above critical resource chain frameworks in order to offer an overview of schedules. Comparing scenario 1 with scenario 2, overall schedule duration is extended, and constraints are massively increased. The CC in scenario 2 contains more activities than CP in scenario 1. Most activities are CC activities except for activity F in scenario 2. Thus, productivity constraints are increased. Additionally, critical resources and, especially, superior critical resources exist everywhere in scenario 1. Contrarily, there is no superior critical resource in scenario 1.

Project planners can meditate on the necessity of considerations to a schedule. The further resource considerations such as resource limitation and productivity options may be required for practical schedules. It can be investigated what happens if the optimized schedule in scenario 2 is tightened: (1) Resources are well utilized; (2) activities are well arranged by considering various productivity options; (3) resource insufficiency can be recognized. However, such a schedule is very sensitive. Any change of schedule elements may affect the whole schedule. Based on the proposed critical resource chain framework, constraints can be recognized. It also can be investigated whether constraints are massively increased from scenario 1 to scenario 2. The sensitive schedule must be kept by observing constraints while performing the schedule. To remind what and where constraints are can be important to schedule performers. Nevertheless, the more considerations when scheduling, the more constraints may be incurred to keep sensitive schedules. The proposed framework is feasible to differentiate constraints for project planners and performers.

Table 1. Project information

Input			
Project indirect cost	$ 1000 / day		
	R1	R2	R3
Daily limit	13 units	10 units	1 units
Cost / unit	$ 200 / day	$ 250 / day	$ 300 / day

Project information						Output: Scenario 1		Output: Scenario 2	
Act.	Successor(s)	Dur.	R1	R2	R3	Start	CP	Start	CC
A	B、C、D、E	3	10	5	1	0	*	0	* #
B	G	6	3	2	0				
		3	5	5	0	3		3	*
C	F、G	12	8	4	0			6	*
		7	12	6	1	3	*		
D	H	20	5	3	0			3	#
		15	7	5	0				
		12	10	8	1	3			
E	--	5	7	3	1	3		18	*
F	I	17	5	2	0				
		15	8	4	0			23	
		13	10	5	0				
		10	13	7	0	10			
G	H	13	5	5	1			23	* #
		11	7	6	1	10	*		
H	I	3	5	3	0			36	* #
		2	8	3	0	21	*		
I	--	2	1	1	1	23	*	39	* #
Overall schedule duration:						25 (days)		41 (days)	
Total resource usage cost:						$ 195,550		$ 174,750	

*: Critical chain 1 #: Critical chain 2

5. Conclusions

This study has proposed a framework of schedule constraints named critical resource chain framework. Schedule constraints are identified and integrated into the proposed critical resource chain framework based on the concepts of TOC, CPM, and RCPSP optimization. Two scenarios of schedules including a CPM based schedule and an RCPSP based schedule with the goal of minimized overall schedule duration are successfully analyzed using the proposed framework. The proposed framework helps planners to understand projects by recognizing constraints. Moreover, further project management issues may be investigated as discussed.

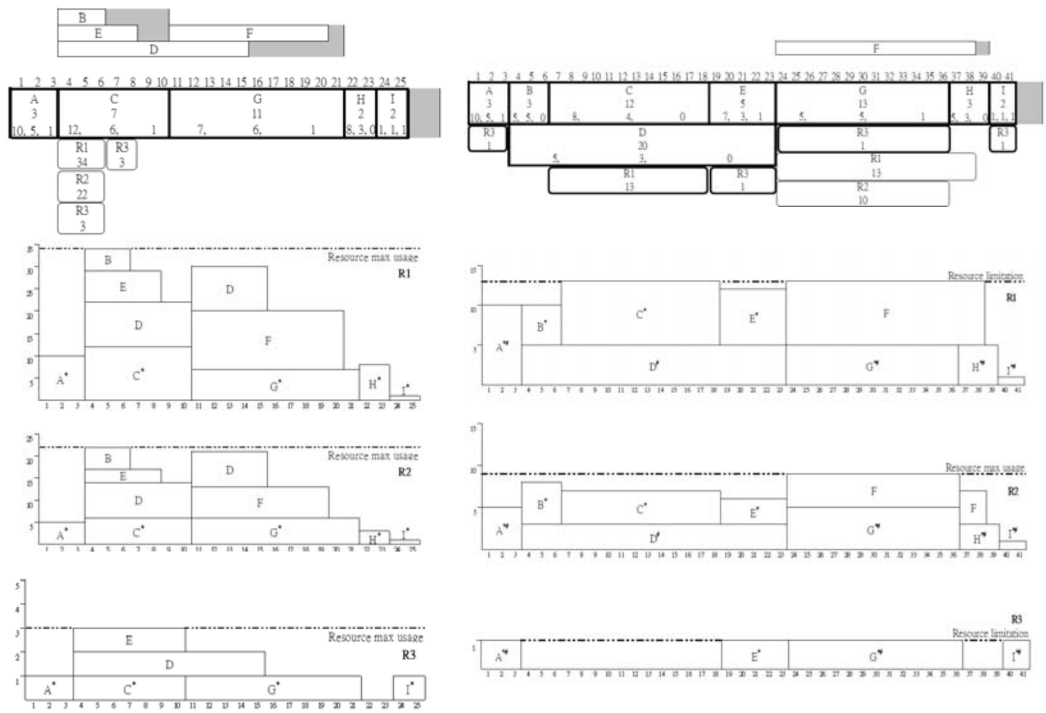


Figure 3. Critical resource chain and resource diagrams of scenario 1 and 2

Acknowledgement

The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financially supporting this research under Contract No. NSC95-2221-E-224-071.

References

[1] Project Management Institute (PMI) (2000). "Project management book of knowledge." Project Management Institute, Drexel, Pa..

[2] Herroelen W., Reyck B.D., and Demeulemeester E. (1998), "Resource-constrained project scheduling: a survey of recent developments." Computers and Operations Research, Elsevier, 25(4), 279-302.

[3] Brucker P., Drexl A., Mohring R., Neumann K., and Pesch E. (1999). "Resource-constrained project scheduling: Notation, classification, models, and methods." European Journal of Operational Research, Elsevier, 112, 3-41.

[4] Goldratt E.M. (1992). "The goal." Great Barrington, MA: The North River Press, 1st ed. 1984, 2nd ed. 1986, 2nd revised ed..

Monitoring and Management of Greenhouse Gas Emissions from Construction Equipment Using Wireless Sensors

Yong-Seon Lee[1], Mirosław J. Skibniewski[1], and Won-Suk Jang[2]

[1]Department of Civil and Environmental Engineering, University of Maryland College Park, MD USA
[2]Building and Fire Research Laboratory, National Institute of Standards and Technology Gaithersburg, MD USA

Abstract

The construction industry alone produced approximately 1.7% of total U.S. greenhouse gas (GHG) emissions in 2002. This is equivalent to 6 percent of the total U.S. industrial-related greenhouse gas emissions. GHG emissions from construction equipments on project sites are highly variable. Standardized method with accuracy and reliability is needed for measuring the GHG emissions from construction equipments. Thus, to comprehend this variation fully and the GHG commodity each construction sites, web-based GHG emissions monitoring system that uses the ZigBee wireless sensors operated by ambient power harvesting for measuring conveniently is needed. According to the results, consuming allocations of each project sites are analyzed. And based on the amount of GHG emissions measured like electricity rates, authorized institute is trading greenhouse gas, regarding the redundant and surplus allocations rates, between project sites. This system would be possible to systematically manage the greenhouse gas and to reduce it eventually. This paper represents the system architecture concept and process for this research.

***Keywords**:* web-based system, greenhouse gas (GHG), construction equipment, wireless sensor, trading center

1. Introduction

The reduction in greenhouse gas emissions, which is partly responsible for global warming, has become a task for maintaining national benefits rather than being simply an environmental protection problem. The recognition of the air pollution problem has changed internationally, as the United Nations Framework Convention on Climate Change (UNFCCC) was adopted through the United Nations Conference on Environment and Development (UNCED) held in Rio de Janeiro in 1992. The terms of the agreement should be fulfilled even by enterprises in their business activities as well as in national matters. Thirty-three states and the District of Columbia in the U.S. regulate automobile emissions by requiring periodic inspections of registered vehicles to ensure that their emissions equipment is functioning properly. However, only one state (New York) and one municipality (New York City) monitor the emissions of heavy vehicles (Clean Diesel Technologies 2008). Fossil fuels used as the power sources by vehicles and machinery, such as those used in construction sites, are largely responsible for the global greenhouse gas problem. In most of existing environmental regulations on enterprise activities, however, certain environmental criteria are designated by the government for enterprises to follow. These methods have some problems in terms of effectiveness, as business people tend to avoid the regulations as much as possible rather than voluntarily observing them. Moreover, as regulations criteria were set without exactly considering the environment of the work site or the characteristics of individual machinery, the industrial circles resisted such criteria. As a result, the development of evaluation technology or measurement technology for air pollution has been delayed. To overcome these disadvantages, it is possible to provide flexibility to users and to further efficiently reduce air environment pollution by introducing the 'Emissions Trading' system of GHG to each construction site. Such systems are now attracting worldwide attention for all industries. By changing the regulations related to air pollution into a system that offers economic benefits and follows market logic. We can influence GHG emissions in construction as well.

To achieve this goal, the first thing to do is to exactly measure amount of greenhouse gases emitted by construction equipment of the construction site. A number of recent research studies have been conducted in order to investigate the GHG emissions from construction equipment working on project sites. Among them, Lewis et al. (2008) investigated the challenges to quantification of GHG emissions from nonroad construction equipments and described associated regulations and incentives for reducing emissions. Rasdorf et al. (2009) researched outline standard procedures for field data collection for construction equipments based on in-use measurement methods. These research studies, however, simply regard measurement and analysis of the greenhouse gas emitted from the construction machinery of the site. Therefore, this study focuses on the measurement of the amount of emissions of greenhouse gas during the duty cycle of heavy vehicles in real time by configuring wireless networks using ZigBee sensors to make accurate measurements of greenhouse gas emissions and then emissions trading of GHG. For convenient measurement, low-priced and stable measurement methods would be carried out through the energy harvesting technology, which can power the wireless sensors by employing vehicle's own vibrations as a source of energy. A more flexible management system will be realized by embodying the greenhouse gas exchange system between the construction sites using the web-based system based on the accurately measured data. Accordingly, it would be possible to systematically manage the greenhouse gas and to reduce air pollution eventually.

2. Wireless sensor

ZigBee is emerging network technology and a wireless communication standard capable of realizing the ubiquitous environment to satisfy requirements. As a superset of IEEE 802.15.4 standard, ZigBee supports the industrial network standards, so that many industrial applications including construction automation, structural health monitoring, and automated control and operation can benefit from the advantages of the technology.

Table 1. Property of the physical layer in IEEE 802.15.4 [ZigBee Alliance, 2005]

Frequency band	2.4 GHz	915 MHz	868 MHz
Number of channels	16	10	1
Bandwidth (kHz)	5,000	2,000	600
Data rate (kbps)	250	40	20
Symbol rate (ksps)	62.5	40	20
Modulation method	O-QPSK*	BPSK**	BPSK
Diffusion method	DSSS***	DSSS	DSSS
Available regions	Worldwide	USA	Europe

O-QPSK* (Offset Quadrature Phase Shift Keying)
BPSK** (Binary Phase Shift Keying)
DSSS*** (Direct Sequence Spread Spectrum)

The ZigBee specification has been released publicly in June 2005, and products supporting the ZigBee standard are widely available in the market. Specified frequency allocations and physical layer recommended by IEEE 802.15.4 is listed in Table 1. ZigBee specification takes advantage of the IEEE 802.15.4 wireless protocols as basic communications method, and expands on this with a robust mesh network, application profiles, interoperability and device description.

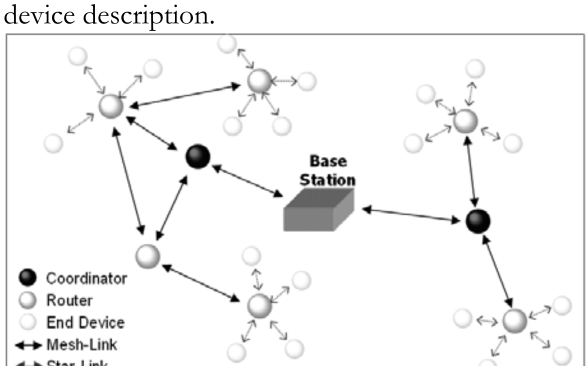


Figure 1. ZigBee Network [Skibniewski and Jang, 2006].

A ZigBee network, shown in Fig. 1, consists of ZigBee coordinators, ZigBee routers and ZigBee end devices [Skibniewski and Jang, 2006]. The coordinator and routers are able to form a star network configuration using PAN coordinator functions, and it is possible to form a multi-hop network by simultaneously configuring the mesh network between the coordinator and routers. The end devices take part in the network communication by linking to the coordinator and routers through star-link networks. The end devices conduct multi-hop communications via connected routers to communicate with other devices connected to the networks. Using the advantages associated with the flexible ad hoc networking, the promise of ZigBee application can be found in robust and reliable, self-configuring and self-healing networks that provide a simple, cost-effective and battery-efficient approach for sensing and network based data communication in construction industry.

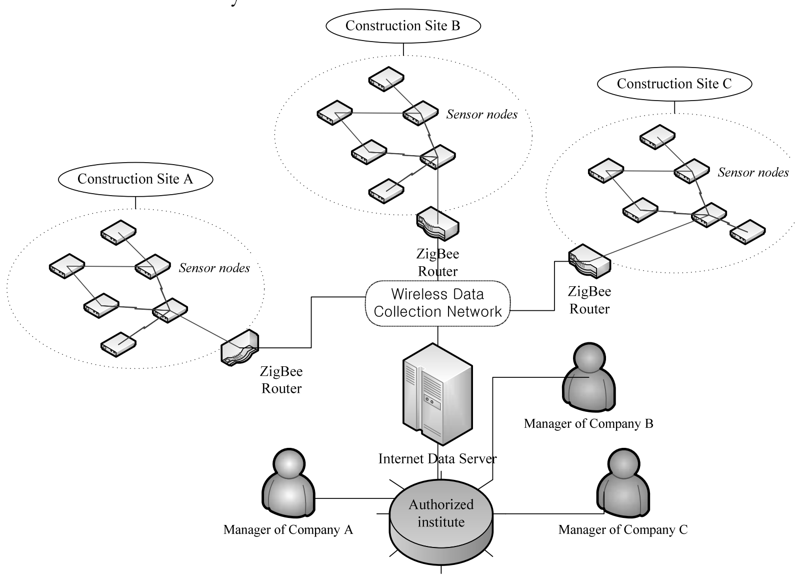


Figure 2. System composition of web-based system

In the wireless sensor network, ZigBee, elimination of undesirable multipath components and fading is an important issue to achieve the localization of the distributed objects. Coordinating is one of the most important tasks in wireless communication networks. To determine the paths and to advertise the identity of a sensor, different coordinating algorithms impose over communication overheads. The sensor nodes with ZigBee protocol transmit the radio signal within the maximum coverage range of 100 m. And this specification of transmission determines the level of power and network topology to be communicated between sensor nodes and ZigBee router.

Using this wireless sensor network, system composition of web-based management system in construction site is made as shown in Fig. 2. ZigBee routers are placed at the location that can cover the entire laydown yard within their trigger ranges to detect the events associated with the movement of distributed sensor nodes. In this network system, measuring data collected to each of the routers is transmitted to the wireless collection network along with the ad hoc path. Different sensor nodes are categorized, identified, and attached to the construction equipments according to the characteristics of greenhouse gas property within the boundary of construction site. The collection network connects to the data server to be managed by authorized institute thorough the internet. The data of greenhouse gas can be used by companies registered through a user interface

3. Energy harvesting for ZigBee sensor

As shown in Fig. 3, when the distributed charge densities applied to the surface of the piezoelectric actuator are converted into voltages as in the following Eqn (1). The variation of the electric potential is obtained on the surface of the piezoelectric actuator; the electrical actuating vectors acting on each factor can be obtained by these actuating voltages (Elvin et al. 2006).

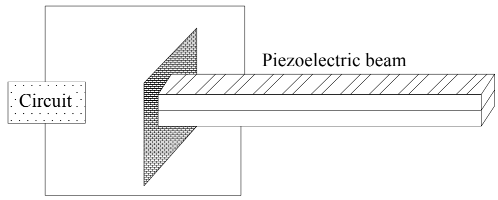


Figure 3. Piezoelectric beam

$$\sigma = \frac{e_z^S}{t_l}\left\{\emptyset_l\left(x,y,z_l-\frac{t_l}{2}\right)-\emptyset_l\left(x,y,z_l+\frac{t_l}{2}\right)\right\} = \frac{e_z^S}{t_l}\cdot Voltage \quad (1)$$

where $\emptyset_l$ presents the $\emptyset_l(x,y,z)$ assumed as function of x,y,z.

Moreover, the variation in electric potential on the surface of the piezoelectric actuator would be obtained by using the Eqn. (2) below.

$$\delta\emptyset_l^t = (z_l - z)\,\delta\emptyset_l^0(x)\big|_{z=z_l-\frac{t_l}{2}} = \frac{t_l}{2}\,\delta\emptyset_l^0(x) \quad (2a)$$

$$\delta\emptyset_l^b = (z_l - z)\,\delta\emptyset_l^0(x)\big|_{z=z_l+\frac{t_l}{2}} = \frac{t_l}{2}\,\delta\emptyset_l^0(x) \quad (2b)$$

where suffixes ‘t’ and ‘b’ show the top and bottom sides of the piezoelectric actuator fixed to the upper side of each beam. Thus, it is possible to configure the electrical actuating vectors expressed as terms of the actuating voltages.

As mentioned previously, this study uses wireless sensors using the energy harvesting technology. Thus, to analyze the validation whether the power of the wireless sensors themselves could be produced with the vehicle’s own vibration, a 3-D model (Doyle 1991) of the moving construction vehicle was developed out as shown in Fig. 4. Likewise, the characteristic of road surface roughness in the construction site was randomly shown in Fig. 5 using a power spectral density function and the roughness coefficient of level poor (Dodds 1973). The equation of road roughness are expressed by choosing the phase angle randomly as in the following. s is the distance ($V \cdot t$).

$$X(s) = \sum_{n=1}^{N}\sqrt{4S_{XX}(\gamma_n)\Delta\gamma}\,\cos(2\pi\gamma_n\, s - \theta_n) \quad (3)$$

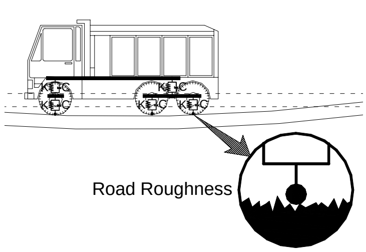


Figure 4. Analytical model

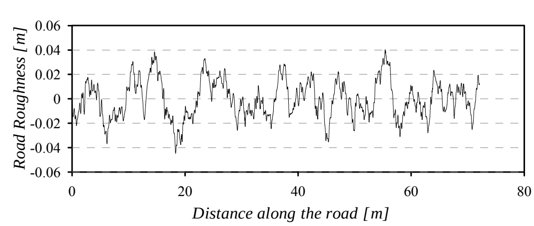
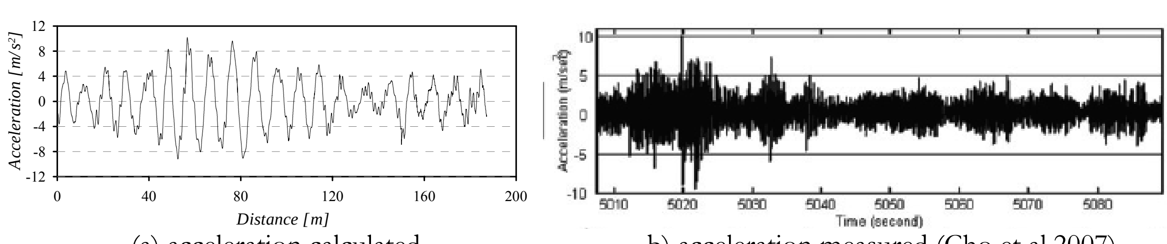


Figure 5. Road roughness



(a) acceleration calculated b) acceleration measured (Cho et al.2007)

Figure 6. Vehicle own responses

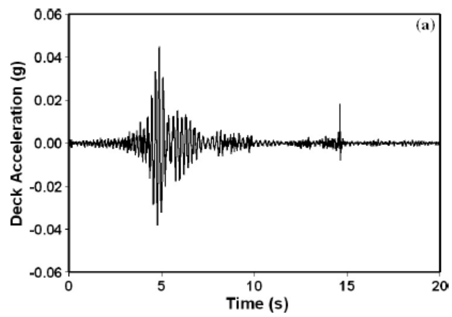


Figure 7. Acceleration of concrete bridge

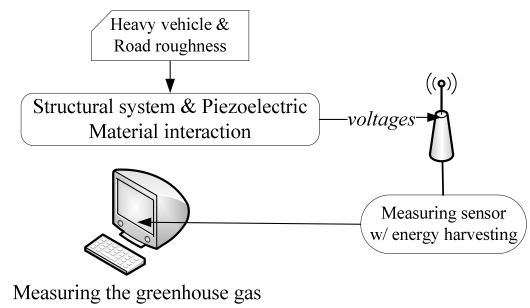


Figure 8. Scheme of energy harvesting

The acceleration response, which was produced from the moving vehicle on the randomly generated road surface roughness, was calculated and real acceleration was measured from heavy vehicle as shown in Fig. 6. Figure 7 shows the acceleration generated by traffic loading on a concrete bridge. This response in Fig. 7 (Elvin et al. 2006) has enough magnitude that can be converted into electrical charge through the piezoelectric material. Therefore, the results were compared with an acceleration value of the vehicle, which was measured from the heavy vehicle running on the road. According to this comparison, it was analyzed that the vehicular vibration itself could be a sufficient energy source for energy harvesting. Figure 8 shows the scheme of energy harvesting using vehicle vibration.

4. Measuring GHG emissions

Currently the overall emission was calculated by determining the amount of GHG emitted from the fuels used as shown in the Eqn. (4). GHG emissions are estimated using the following equation (Mui, et al. 2007)

$$Emissions = \left[\frac{Gallons}{Mile}\right]\left[\frac{Carbon}{Gallon}\right]\left[\begin{matrix}Vehicle\ Miles\\ Traveled\end{matrix}\right] \quad (4a)$$

$$E = F\ (Fuel\ Consumption) \times C\ (Carbon\ Contour) \times A\ (Activity) \quad (4b)$$

Each of the three factors in the equation-per-mile fuel consumption (F), fuel carbon content (C) and VMT (vehicle-miles traveled) (A)-contributes to overall emissions. Thus, reductions in one parameter may be offset by increases in another. Comprehensive GHG policies will consider all three factors. In the other method, a simple equipment called portable emissions measurement systems (PEMS) was installed at the outlet of automobile exhaust fumes to measure the vehicle's emission amount of GHG. This method, however, cannot perform a long term calculation of the emission amount of GHG and have some restrictions on installation and operation under various environments due to problems such as equipment size, installation methods, and acquisition of power source. So, this method is being used mostly in laboratory measurement. To figure out these restrictions, this research tries to measure GHG from equipments used in the construction site by employing the wireless sensors. The testing procedure used in this study is shown in Fig. 9. First, the measurement ranges in each construction site was selected, and data, including various types of vehicles to be used in these construction site, working hours, and idle hours were preliminarily investigated and put in order. Next, main GHG emission sources of each construction area were determined, and these sources were measured to determine reference emission amounts per hour.

In addition, this study performs uncertainty interpretation of the measured GHG data. It analyzes the uncertainty of data caused by the deficiency of accuracy in the measurement values using the wireless sensors. In the common case, a result, R, to be function of n measured variables depends on measurements whose values are $x_1 \pm w_1$, $x_2 \pm w_2$, ….., $x_n \pm w_n$.

$$R = f(x_1, x_2, x_3 \ldots\ldots, x_n) \quad (5)$$

w_n is the uncertainty in the variable. So maximum uncertainty in R can be estimated using the Eqn. (6).

$$w_R = \sqrt{\sum_{i=1}^{n}\left(w_i \frac{\partial R}{\partial x_i}\right)^2} \tag{6}$$

This equation is known as square root mean of the sum of the squares (RSS). The confidence level in the result will be the same as the confidence levels of the uncertainties in the measured variables. Such an analysis can measure the GHG allocation amounts per scale of each construction site, and also calculate the exact measurement values to enable transactions between the construction sites.

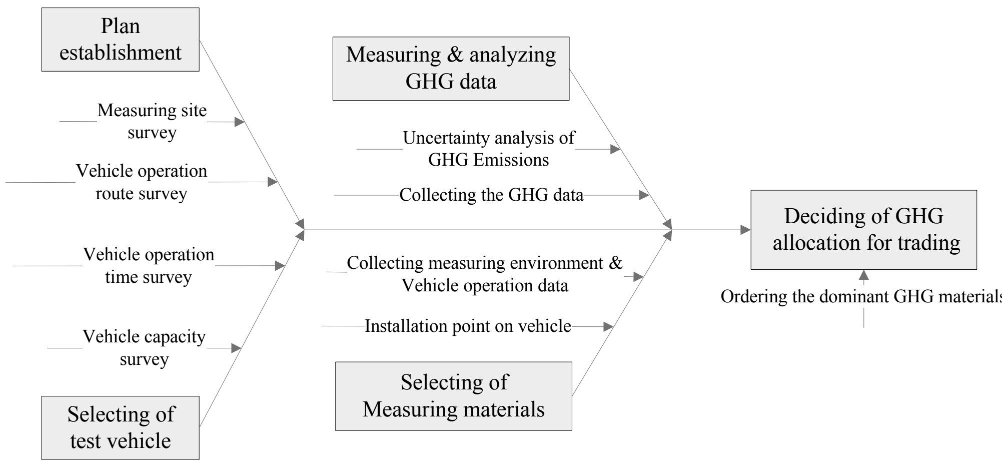


Figure 9. Experiment procedure

5. Web-based GHG trading schemes for construction sites

It is possible to determine the exact GHG emission amounts of construction equipment used in the construction site by using self-powered wireless sensors with the aforementioned piezoelectric material. Based on these exact emission amounts, an approved institution intermediates transactions between a construction site having surplus emission amounts of GHG and a construction site having excessive emission amounts. The intermediation method is to send the measurement values to the Internet, thereby constructing a web-based management system for managing these measurement values. The advantage of this a management system is the ability to check the emission amounts of GHG in real time in each construction area, and to enable one construction area, which needs additional allocations of GHG due to the excessive emission amounts, to confirm the allocations of another construction area that has surplus allocations and to immediately trade in with this construction area. Besides, the profits obtained through such a GHG trading system would be reinvested as research subsidies for the development of new energy sources, the construction of new generation plants, and repair and extension works of the degraded power plants. The flow of this trading system is shown in Fig. 10.

To establish this GHG trading, web-based management system need to develop approach system to be accessible data from the computer connected to the gateway as well as from anywhere. Software written in PHP to access the database and display results needs, therefore, to access the sensor database over the Internet. It lets the user modify the database to add more sensor nodes or sensor boards and allows him to query the database to find necessary data. The important parts of this web system include access to the data, information about each node in the network. The site with a familiar web environment allows the user to both read data that have been recorded and to modify the sensor network by adding descriptions of nodes or sensor boards.

Another function to be implemented in the website is the ability to change the information on the sensor network when nodes are moved to a different location and a specific node number is no longer needed, or when a new node is introduced into the network. Data are easily accessed via the website, and sensor nodes could be added to the network with little trouble. By breaking down the monitoring system into data acquisition, data storage, and data retrieval components, a flexible system is created that can be modified by users to meet their needs while allowing for advances in wireless sensing and networking technologies to augment monitoring capabilities without disrupting the operation of the entire system. The benefits of a system such as the one described in this work lie in the ease with which engineers can obtain information from construction equipment.

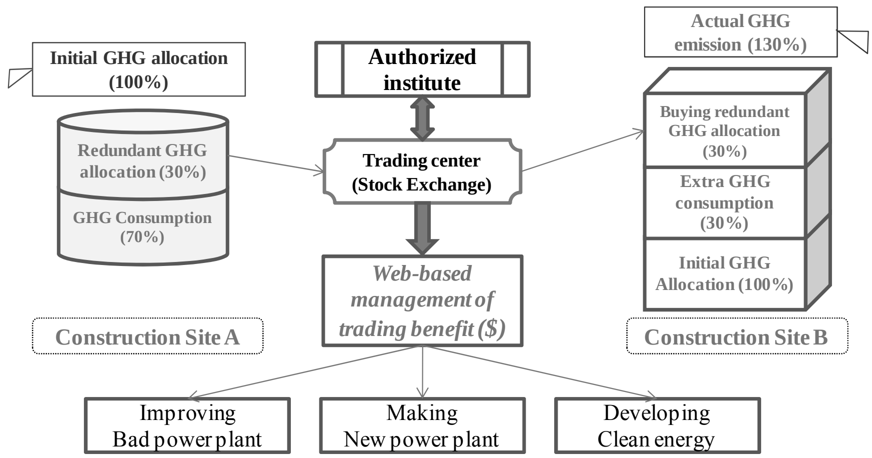


Figure 10. Diagram of web-based GHG trading

The committee of authorized institute would be responsible for sponsoring data integrity and reporting standards. Specifically, they would need to determine platform specific encryption standards and reporting pathways. Develop a select list of approved vendors for direct GHG monitoring and on-board computer interface systems. It would also be important for the committee to determine what information is important and how it should be analyze to help bolster the greenhouse gas emissions trading system. Furthermore, to enhance the reduction of greenhouse gases needs kinds of incentives. In a sense, it is necessary to design and implement a pathway for enhancing greenhouse gas trading of project sites with construction equipment within the international GHG emission trading framework.

6. Conclusions

This ongoing study aims to formulate a management solution for GHG trading between the construction sites based on the measured emission amounts. In order to analyze if vehicular vibration is sufficient to converts into a power source of the wireless sensors by using piezoelectric material, a vehicular response is compared with the substantial measurement value by considering the vehicular model and road surface condition of the construction field,. As a result, it was analyzed that the vehicular vibration was enough to be changed into the power source, which makes it possible to produce power using the vehicular vibration. Furthermore, the exact long term GHG emission amounts could be measured using the self-powered wireless sensors, and GHG trading between construction sites would be possible through the web-based management system on the basis of the exact measurement amounts of GHG. Likewise, the profits made by greenhouse gas trading would be invested for the improvement of degraded power plants, in the construction of new relatively “clean” plants with none to minimal GHG emission amounts, and the development of new clean energy. Accordingly, greenhouse gas emissions, which currently pose as a huge global problem, can be eventually reduced.

References

[1] Cho, B.H. and Lee, S.H. (2007). “Vibration performance of modular units under transportation.” Journal of RIST Research, Research Institute of Industrial Science & Technology, 21(3), 213-221.
[2] Clean Diesel Technologies. [Internet]. (2008). Clean Diesel Technologies Reduces Emissions for Skanska on New York City's Largest Public Works Project.
[3] Dodds, C.J. and Robson, J.D. (1973), “The Description of Road Surface Roughness.” Journal of Sound and Vibration, Elsevier, 31(12), 175-183.

[4] Doyle, J.F. (1991). "Static and Dynamic Analysis of Structures: With an Emphasis on Mechanics and Computer Matrix Methods." Springer.
[5] Elvin, N.G., Lajnef, N., and Elvin, N.G. (2006). "Feasibility of structural monitoring with vibration powered sensors, Smart Materials and Structures." Institute of Physics Publishing, 15, 977-986
[6] Lee, Y.-S., Kim, S.-H., and Jang, W.-S. (2005). "Analysis of Sound Radiation from the Coupling Effect of Vibrating Noise and Moving-Vehicles Noise on Bridges." Canadian Journal of Civil Engineering, NRC Research Press, 32(5), 881-898.
[7] Lewis, P., Rasdof, W., Frey, H.C., Pang, S.-H., and Kim, K. (2008). "Requirements and incentives for reducing construction vehicle emissions and comparison of non-road diesel engine emissions sources." Journal of Construction Engineering and Management, American Society of Civil Engineers, in press
[8] Mui, S., Alson, J., Ellies, B. and Gauss, D. (2007). "A Wedge Analysis of the U.S. Transportation Sector." Environmental Protection Agency. EPA 420-R-07-007, 23 pp.
[9] Rasdof, W. Frey, H.C., Lewis, P., Kim, K., and Pang, S-H., (2009). "Field Procedures for Real-World Measurement of Emissions from Diesel Construction Vehicles." Journal of Infrastructure System, American Society of Civil Engineers, Submitted
[10] Skibniewski, M. J. and Jang, W.-S. 2006. "Ubiquitous computing: object tracking and monitoring in construction processes utilizing ZigBee networks." in Proc of International Symposium on Automation and Robotics in Construction 2006, ISARC 2006, 287–292.
[11] ZigBee Alliance. ZigBee specification. ZigBee Document 05347r06, San Ramon, CA, USA, June 27, 2005.

Light-weight 3D LADAR System for Construction Robotic Operations

Yong K. Cho[1] and Diego Martinez[2]

[1]Assistant Professor (corresponding author), Construction Engineering and Management Division, Charles Durham School of Architectural Engineering and Construction, College of Engineering, University of Nebraska-Lincoln, 1110 S. 67th St. Peter Kiewit Institute (PKI) 104C, Omaha, NE 69182; PH(402)554-3277;FAX (402)554- 3850; email: ycho2@unl.edu

[2]Graduate Research Assistant, Computer Engineering and Electronics Department, College of Engineering, University of Nebraska-Lincoln, 1110 S. 67th St. Peter Kiewit Institute (PKI) 118A, Omaha, NE 69182; PH(402)554-3277;FAX (402)554- 3850; email: dmartinez@mail.unomaha.edu

Abstract

This paper presents an on-going research on a light-weight 3-dimensional (3D) laser distance and ranging (LADAR) system development at the Peter Kiewit Institute (PKI) in University of Nebraska–Lincoln. The developed LADAR can be readily applied to several construction applications related to automated or robotic tasks, mobile robot navigation, reverse engineering, quality assurance/control, schedule control, and safety. A design concept of a LADAR system on a dynamic mobile platform is introduced which is especially designed for precise construction robotic operations such as welding, bolting, and connecting materials. Barriers for rapid 3D graphical workspace visualization at construction sites are discussed as well. The preliminary laboratory test results demonstrate that the 3D LADAR scanner developed in this study provides reliable scanned data. Unlike most of the heavy and bulky commercial LADAR products, the developed economical LADAR system can be used on mobile platforms and manipulators or confined area scanning applications due to its small size and light-weight characteristics.

Keywords: 3D scanner, LADAR, laser, visualization, robotics, construction automation, error

1 Introduction

In construction, significant time is spent every day on job sites to identify and track construction materials and equipment and to capture, align and compare field measured data to planned data to detect defects and improve quality control. This requires rapid recognition and accurate measurement of objects in the field so that timely, on-site decisions can be made. Especially 3-dimensional (3D) graphical visualization of the site can help to optimize material tracking and automated equipment control, significantly improve safety, and enhance a remote operator's spatial perception of the workspace. FIATECH (2008) envisions that construction sites will become more 'intelligent and integrated' as materials and equipment, and people become elements of a fully sensed and monitored work environment.

Although studies in several fields have proven that 3D visualization of the work environment can significantly enhance construction defect control, material tracking, schedule control, and automated equipment control, unstructured work areas like construction sites are difficult to visualize graphically because they contain highly unpredictable activities and change rapidly. Automated construction site operation, e.g., robotic operation, requires real-time or near real-time information about the surrounding work environment, which further complicates graphical modeling and updating. Especially, solid material handling tasks in construction require not only rapid visualization of unstructured workspace but highly accurate position data for safe and secure physical contact between a target object and an end-effector of automated mobile equipment or robots.

This paper discusses the current 3D visualization technologies and their limitations for construction applications and introduces a 3D LADAR scanner being developed for robotic applications in construction sites.

2 Current 3D Visualization Technologies

This section discusses applicability of current 3D visualization technologies at construction sites.

2.1 Machine Vision systems

CCD camera-based machine vision systems are useful to visualize a workspace under well controlled environment. However, they may not perform very well to get an accurate pose of the target objects for heavy construction applications which require long distance position measurements. Furthermore, as with any vision application, consistent lighting is critical to obtain accurate workspace information. It explains why most of the robotic applications with a vision system are mainly limited to indoor applications, including automotive, electronics, packaging, machining, and food. Getting high position data accuracy using a vision system at unstructured indoor and outdoor construction sites is impractical due to extreme or inconsistent daylight conditions, bad weather, and clouds of dust. In that case, a distance sensor, such as a laser or an ultrasonic sensor, might better enable the robot to draw a bead on the target (Sprovieri, 2007).

2.2 Flash LADAR Research

An emerging range sensing technology, called 3D range camera or Flash LADAR, is based on the time-of-flight measurement principle which uses light. The estimated accuracy in Z axis with one pixel of the Flash LADAR (SR4000) is +/- 1cm with up to 54 frame rate (FPS) under well controlled indoor lighting conditions. SR4000's measurement distance ranges from 0.3 to 5 meters (Mesa Imaging, 2008). While LADAR/LIDARs need to scan objects or scenes to collect point data sets (point clouds), flash LADAR systems do not require a scanning pointing mechanism. Compared to commercially available LADAR/LIDARs, the advantages of using a flash LADAR are that it is smaller, inexpensive, and forms 3D images in real time. The disadvantages include a limited view size and lower accuracy and resolution. In addition, SR4000, the latest version of Flash LADAR, is designed mainly for indoor applications. The noise level makes it impossible to work in direct sunlight.

2.3 ZScanner

Z Corporation (2008) has developed the first hand-held self-positioning scanner on the market. This digital scanner uses the subject part being scanned to establish its spatial reference eliminating the need for fixed-position tripods, bulky mechanical arms or external positioning devices that make hard-to-reach surfaces to scan. Uniquely part-referenced, the system allows the target object's moving during scanning and a real-time image of the surface being scanned can be seen. It captures data in one continuous scan rather than in numerous shots from fixed positions, eliminating hours of post-processing time. Scanning resolution and accuracy are within 40μm (microns) and detects changes in surface height down to 50μm. The scanner supports STL and TXT file formats. To position itself, the ZScanner uses reflective targets, which are applied to the surface of the part to be scanned and/or the area adjacent to the part. These reflective targets can be quickly and randomly applied. During the scanning process, the ZScanner locates and captures the reflective positioning targets in real time. Their respective positions are calculated in reference to the scanner and then recorded. The targets on the object create patterns recognizable by the ZScanner. The patterns defined by the positioning features do not repeat because the targets are applied randomly. This pattern recognition allows the scanner to position itself in the same way that GPS devices use known satellites to establish their position on Earth (Z Corporation, 2008).

In this study, the ZScanner is tested for scanning a fake duck model on which many reflective dots (targets) are attached. The scanner models a 3D graphical duck in real time while it is scanning a duck model (Figure 1).

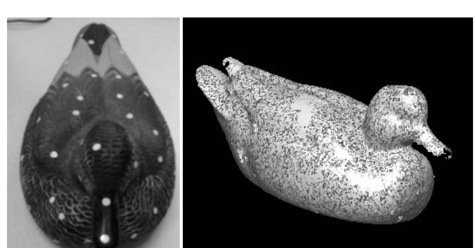


Figure 1. Example duck model scanned with ZScanner

The limitation of this scanner is its short measurement range (about less than 1 meter). Also the scanner cannot model a complex surface well like a human ear if it cannot have a clear shot from the both cameras. In this case, the camera sees just with one "eye" and not with the other, which models an incomplete shape. The example of this problem is shown in Figure 2.

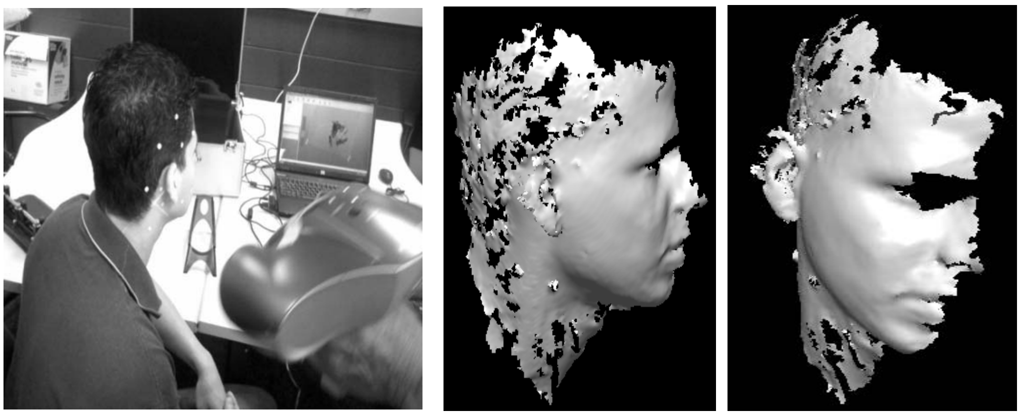


Figure 2. Example of incomplete model with complex surfaces, e.g., ear

2. 4 LADAR/LIDAR (3D Scanning) Technologies

Laser distance and ranging (LADAR) or light detection and ranging (LIDAR) technology has been used to create 3D as-built models of structures and scenes for quality control, surveying, mapping, reverse engineering, terrain characterization, autonomous vehicle navigation, and vehicle-based safety and warning systems (Cheok et. al. 2005). Time-of-flight (TOF) or phase shift measurement is generally used in these laser scanning systems which produce a data set that consists of multiple X,Y,Z, and I values, where I represents the intensity of the reflected laser return. The points in a data set are referred to as a point cloud. State-of-the art LADAR technology, e.g., Leica ScanStation and RIEGL LMS-Z390i, provides about 4mm distance accuracy and 6mm position accuracy up to 50 meters for a single measurement. The error becomes greater when point clouds are obtained from a scene. Accuracy depends on a number of additional factors beyond the underlying sensor accuracy, including measurement goal type, object size, surface orientation, and even surface material, e.g, darkness, reflectivity (Akinci et al. 2006).

2.4.1 Time-of-Flight (or Pursed) vs. Phase Shift

The time-of-flight scanner has a capture rate of around 4,000-6,000 points per second and is able to capture between 130-200m diameter of data depending on type and reflectivity of the material/object being scanned. Unlike the time-of-flight scanner, the phase shift scanner's data capture rate is very fast, e.g., up to 625,000 points per second for Z+F laser scanner™, thus it produces high point cloud density in a short period. However, the phase shift scanner acquires shorter distance measurement (around 40m-80m). Leica ScanStation2™ and RIEGL™ are good examples of time-of-flight scanners; and Faro™ and Leica HDS 6000™ are good examples of phase shift scanners.

3. 3D LADAR Prototype development

LADAR/LIDAR system is superior to machine vision and flash LADAR under aforementioned unfavorable working condition. However, LADAR/LIDAR technology still has limitations to be applied to some automated construction applications. Most of the LADAR commercial-off-the-shelf products are difficult to apply as they are for the dynamic robotic operation in construction without reducing substantial amount of size and weight. They are too heavy and expensive to be mounted on dynamic robotic systems. For example, Leica's Scanstation with a battery pack weighs about 30 kgs. To resolve this issue, this research develops a small and light LADAR system consisting of a 2D line laser and a pan and tilt unit (PTU) which can be mounted on mobile platforms or used confined area scanning applications from an articulated robotic arm.

The 2D line laser as a main component for the LADAR system is a precise spinning mirror assembly that deflects a laser beam 90 degrees, sweeping it through a full 2D circle as the mirror rotates. When coupled with a pan and tilt unit, then, the 2D line laser creates a light-weight and economical 3D scanner. It is important to consider key factors when selecting a line laser according to application goal type. The key factors include resolution, accuracy, repeatability, distance range, and indoor or outdoor use.

In this research, the custom-built LADAR scanner provides point clouds up to 200K points per second from a scene within 8mm accuracy for about 15m distance. The design concept of the 3D LADAR scanner developed in this research can be applied to several different applications by replacing a line laser to meet application's requirement. For example, industrial welding/ bolting robots and medical scanning applications may need high resolution, accuracy, and repeatability for indoor use. A short distance range may be acceptable. The applications for cranes and construction site safety primarily need a long-range line laser designed for outdoor conditions.

To have target-focused measurements from a scene, a single-axis laser rangefinder is coupled with the 3D LADAR system. For the convenience of laboratory experiment and software development, the whole sensor and control units are mounted on a custom-built mobile cart which can be also simulated as a mobile construction platform. To scan objects which are in blind spots, or need multiple surface scans, the 3D LADAR system is mounted on a rotating steel arm controlled by an accurate actuator which can simulate a robot arm. Based on the current laser mounting configuration, multiple Degree-Of-Freedom (DOF) kinematics is solved to obtain x-y-z point values for both the LADAR and a laser rangefinder systems (Figure 3). In analyzing the kinematics for the LADAR system, the Denavit-Hartenberg (D-H) parameters (Craig, 1986) is applied. The D-H parameters describe the positions of links and joints unambiguously. Each link and joint pair can be described as a coordinate transformation from the previous coordinates system to the next coordinate system. Figure 4 shows the point clouds of the scanned room area with the developed 3D LADAR system.

3.1 Noise reduction

Parameters affecting the noise level are sampling rate, ambient light and temperature, surface orientation of the target, target surface properties including color, i.e., gray level, and material types, mixed pixels, and poor visibility conditions (Akinci et al. 2006). Ambient light and temperature did not significantly affect the range accuracy but the accuracy was turned out to be very sensitive to surface orientation, reflectivity and color (gray level) in this research. Based on the target gray level, an amplitude rate needs to be accordingly adjusted to get better results. The amplitude rate determines the intensity or brightness of light. The test results show that the amplitude rate significantly affects the noisy level when it was adjusted in advance according to target's surface color. Figure 5 shows the test results using the KUKA testbed. The KUKA robot is used to simulate welding and pick and place tasks based on the 3D scanned workspace information to evaluate the accuracy and efficiency of the developed LADAR system.

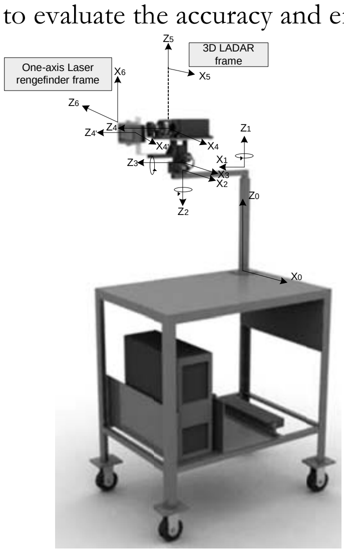


Figure 3. Prototype LADAR system on a cart

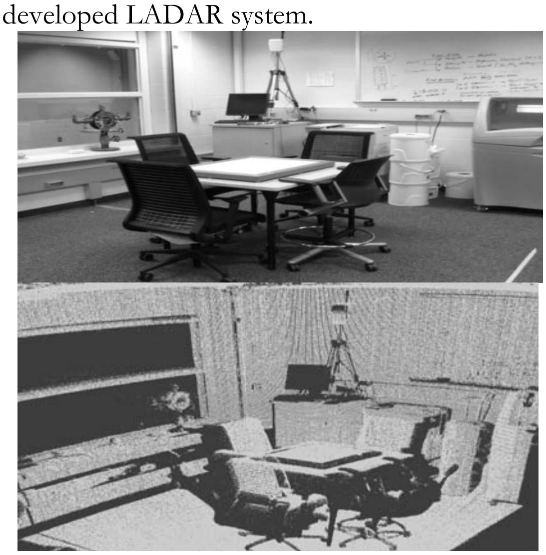


Figure 4. Photo image (top) and scanned point clouds image (bottom) for a lab meeting area

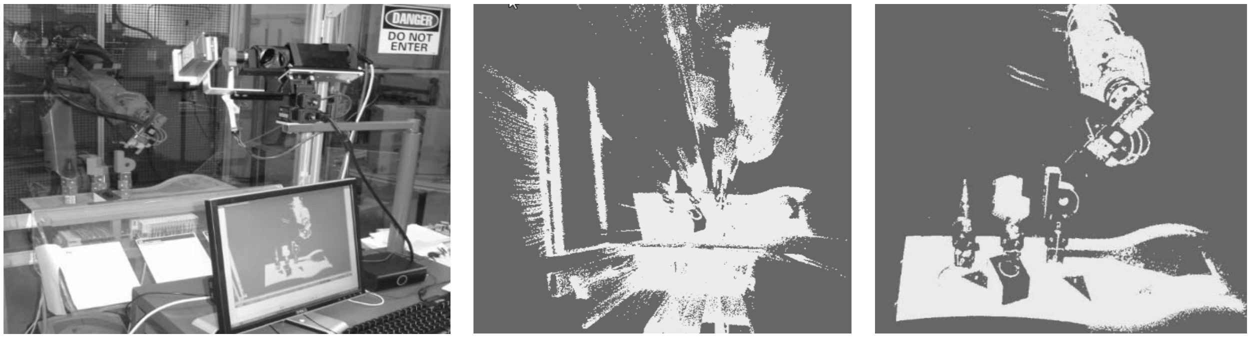


Figure 5. (a) KUKA robot testbed with targets (b) Noisy LADAR point clouds scanned without range guidance & without amplitude adjustment (c) Target-focused point clouds and reduced noises with amplitude adjustment

3.2 Current Research Efforts

As an on-going research, the research team currently focuses on the following two tasks:

3.2.1 Dynamic multi-scans from a robotic arm

Two or more multi-scans from different views can be theoretically achieved when a scanner is mounted on a robotic arm and the scanned scenes can be merged while scanning without losing track of the scanner's position in reference to the target. In reality, however, a robot arm or manipulator generates position errors due to accumulated feedback errors of the rotary and prismatic joint sensors, and lost actuator motion due to backlash. Presence of payloads, kinematic and dynamic states influence error attributes as well (Cho et al. 2004). The research team currently investigates how error attributes created from rotary actuators used for rotating arm and the pan and tilt unit influence theoretical kinematics solutions. Then, the identified error attributes are used to calibrate the LADAR position estimation.

3.2.2 Automatic target detection and registration

Recognizing objects by establishing point correspondences between a CAD model and a scene data set is not a new idea (Grimson, 1990). However, it is still not an easy task especially when the scene data set contains a large number of noisy point measurements. It is also a challenging task if a solid object needs to be automatically registered to that noisy scene data set. Currently, this registration process is manually done. This research is currently developing an algorithm which can estimate and predict surfaces from the noisy point measurements by rearranging the points for more effective target detection and registration processes. Figure 6 briefly illustrates this process.

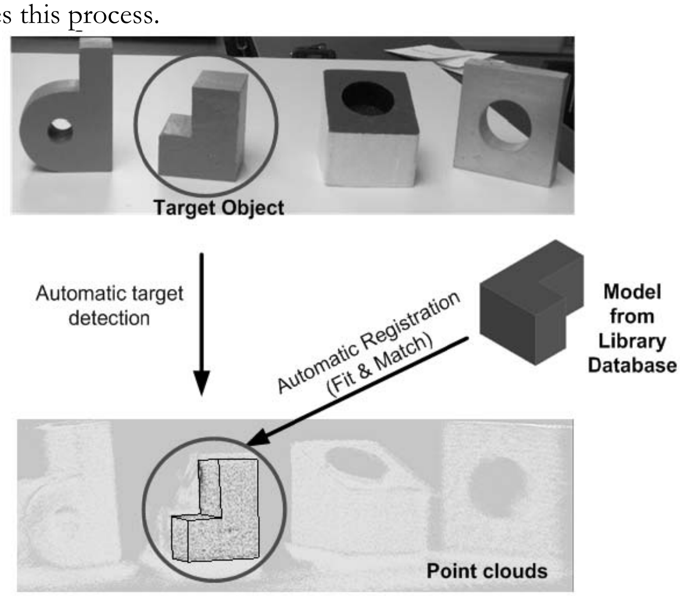


Figure 6. Automatic target detection and registration process

While the research team is considering several different approaches to register an object into the 3D space, the LADAR image processing method is introduced here.

Edge detection

To automate the target object recognition, an intensity LADAR image is measured and analyzed to automatically find target object's boundary lines (Figure 7).

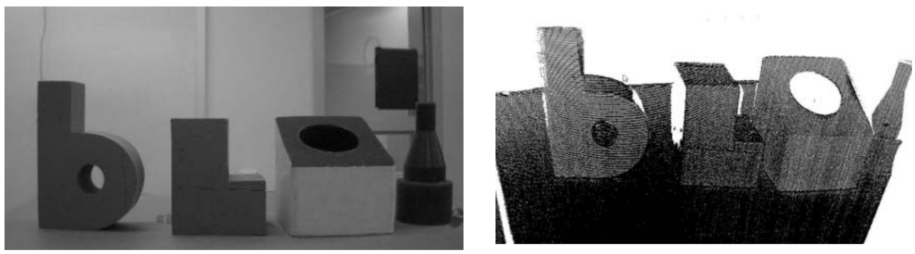


Figure 7. Target objects (left) and binary edge detection of scanned point cloud image (right).

3D model registration

To fit a 3D CAD model to the LADAR image, the model-based pose iteration process using the Iterative Inverse Perspective Matching algorithm developed by Wunsch and Hirzinger is being considered (1996). As there is no distance metric relating 3D point coordinates to 2D image coordinates, the 3D model will be projected into the LADAR image plane and to match features in image coordinates. The 2D to 3D registration problem can be formulated as follows (Wunsch et al. 1996):

$$E(R, t) = \sum_{p \in P} \left(d\left(CLP\ (p, RX + t) \right) \right)^2$$

Where, R is a 3x3 rotation matrix and t is a translation vector.

$$E(\boldsymbol{R}, \boldsymbol{t}) = \sum_{i=0}^{N} \varrho(yi - (Rxi + t))$$

$$\varrho(x) = \begin{cases} 0.5x^2, & \text{if } |x| < a \\ a|x| - 0.5a^2, & \text{otherwise} \end{cases}$$

Combing the computation of closest points with a robust pose estimator leads to the following iterative registration procedure:

for k=0,1, …… **do**
  $[x^k, y^k] = \mathrm{CLP}(P, X^k)$
  [R,t]= register (x^k,y^k, w^k)
  X^{k+1}= R X^t + t
  W^t= ComputeWeight
endfor

Figure 8 illustrates the fitting and matching process using Iterative Inverse Perspective Matching algorithm (Chen and Medioni, 1992; Besl and McKay, 1992; and Wunsch and Hirzinger,1996).

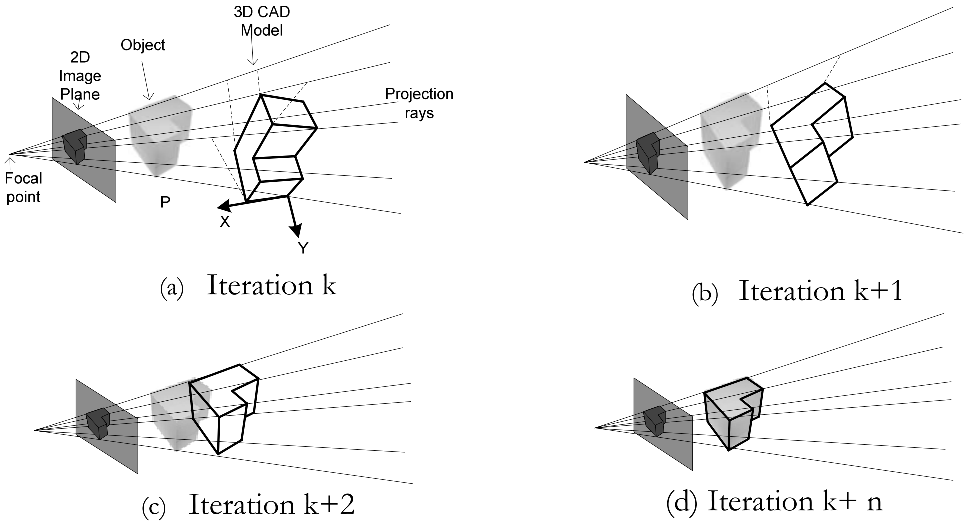


Figure 8. Establishing correspondences between an image and a 3D model based on closest points.

4. Conclusions

A 3D LADAR scanner development research is introduced in this paper. The developed small and light-weight scanner is readily applicable for several dynamic construction automation applications such as robotic manipulator control, robotic inspection, crane application, road profiling, and confined space scanning, where many of commercial scanners have not been well utilized because of their bulky and heavy mechanisms, and expensive purchasing cost. As a major contribution of this research, such a small LADAR system developed in this study can be a useful component to make more "intelligent and integrated" construction sites by providing a highly accurate rapid 3D workspace for automated construction equipment or robotic operations. As an on-going research, the research team continues to conduct the tasks including dynamic multi-scan and registration, rapid data process for real-time robotic operations, error modeling, 3D model registration algorithm development, and mobilization for construction site navigation.

References

[1] Akinci, B., Boukamp, B., Gordon, C., Huber, D., Lyons, C., and Park, K. (2006). "A Formalism for Utilization of Sensor Systems and Integrated Project Models for Active Construction Quality Control," Automation in Construction, Vol. 15, No. 2, February, pp. 124-138.
[2] Besl, P., and McKay, D. (1992) "A method for registration of 3-D shapes," IEEE Trans. Pattern Analysis and Machine Intelligence, 14(2):239-256, February.
[3] Chen, Y. and Medioni, G. (1992). "Object Modeling by Registration of Multiple Range Images." Image and Vision Computing, 10(3), 145-155.
[4] Cho, Y., Haas, C., Sreenivasan, S., and Liapi, K. (2004). "Error Analysis and Correction for Large Scale Manipulators in Construction," the ASCE Journal of Construction Engineering and Management Vol. 130, (1), pp.50-58.
[5] Craig, J. (1986). Introduction to Robotics: Mechanics & Control. Addison-Wesley.
[6] FIATECH (2008). "Capital Projects Technology Roadmap: Element 4 Tactical Plan, Intelligent and Automated Construction Job Site (IACJS), <www.fiatech.org/projects/roadmap/cptri.htm> (Dec.11, 2008).
[7] Gilsinn, D., Cheok, G., Witzgall, C., Lytle, A. (2005), "Construction Object Identification from LADAR scans: An Experimental Study Using I-beams," NIST Integragency Report.
[8] Grimson, W. (1990). Object recognition by computer: The Role of Geometric Constraints, MIT Press, Cambridge, MA, 1990.
[9] Grimson, W. (1990). Object recognition by computer: The Role of Geometric Constraints, MIT Press, Cambridge, MA, 1990.

[10] Johnson, A. and Hebert, M. (1998). Surface matching for object recognition in complex three-dimensional scenes. Image and vision computing : Projection-based transforms, vol.16, n9-10(20 ref.),pp.635-651.
[11] Mesa Imaging (2008). SwissRanger™ SR4000, <http://www.mesa-imaging.ch/prodview4k.php>, (Dec. 8, 2008).
[12] Sprovieri, J. (2007). "Vision Systems Help Robots Find, Assemble Parts," Assembly Magazine, issued on April 1, 2004, www.assemblymag.com/CDA/Archives/bc0975c53f5c9010VgnVCM100000f932a8c0, (September 2007).
[13] Wunsch, P. and Hirzinger, G. (1996). "Registration of CAD-models to images by iterative inverse perspective matching," in Proc. 13th Int. Conf. on Pattern Recognition, Vienna, Austria, pp. 78-83.
[14] Z Corporation (2008). ZScanner, <http://zscanner.com >, (Dec. 8, 2008).
[15] Zheng, Q., Der, S., and Chellappa, R. (2001). "Model-based Target Recognition in Pulsed Ladar Imagery," IEEE Transaction on Image processing, vol. 10, No.4, April, pp. 565-572.

Experimental evaluation of a robotic bolting device in steel beam assembly

Kyungmo Jung[1], Youngsu Chu[1], Baeksuk Chu[2], Daehie Hong[2#],
Shinsuk Park[2], Myo-Taeg Lim[3], Yongkwun Lee[4], Sung-Uk Lee[5],
Kang Ho Ko[6] and Min Chul Kim[7]

[1]Graduate School, Korea University, 5-1, Anam-dong, Sungbuk-gu, Seoul, 136-701, Korea
[2]Division of Mechanical Engineering, Korea University, 5-1, Anam-dong, Sungbuk-gu, Seoul, 136-701, Korea
[3] School of Electrical Engineering, Korea University, 5-1, Anam-dong, Sungbuk-gu, Seoul, 136-701, Korea
[4]Center for Cognitive Robotics Research, Korea Institute of Science and Technology, Seoul, Korea
[5]Nuclear Robotics Laboratory, Korea Atomic Energy Research Institute, Daejeon, Korea
[6]R&D Center mPLUS Corp., Hwaseong, Gyeonggi, Korea
[7]R&D Center ROBOTOUS Inc., Seongnam, Gyeonggi, Korea
[#] Corresponding author. Tel.: +82 2 3290 3369, Fax.: +82 2 926 9290, E-mail address: dhhong@korea.ac.kr

Abstract

This paper deals with a robotic system for the steel beam assembly task in building construction. Applying the robotic system on the steel beam assembly, we pursue such advantage as improving safety for workers and reducing construction period and cost. The robotic bolting device suggested in this paper has a gantry-type moving apparatus which includes a bolting end-effector. The bolting end-effector is designed especially for bolting mechanism using TS(Torque Shear) type bolt. In order to perform the bolting assembly tests, we made an experimental testbed which was similar to a real steel construction environment. Through the tests, we acquired feasible results for applying the robotic bolting device to real steel construction sites.

Keywords: Construction robot, bolting, steel beam assembly, robotic bolting device.

1. Introduction

In recent years, various studies have been attempted to employ robotic technologies to construction field. As an example, in Korea, a research on the development of the Construction Automation System of High-rising Building was initiated[1,2]. This system is based on CF(Construction Factory) technology[3,4]. CF is a factory-like structure in which several construction robots are equipped to automate construction works. This project consists of four research categories like follows: (1) Construction automation system planning and integration for robotic crane based high-rise building structure, (2) Climbing hydraulic robot and CF structure technology, (3) Robotic crane based construction material installation technology, and (4) RFID(Radio Frequency IDentification) and multi-DOF CAD based intelligent construction material supply system. Figure 1 shows a 1/20 size model of the construction automation system.

This paper focuses on "(3) Robotic crane based construction material installation technology". In order to perform the task (3), an automation system for robotic beam assembly was suggested[5]. In this study, the robotic bolting device for beam assembly was briefly reviewed and extensive experiments were performed to obtain bolting criterion to assure successful bolting tasks. With the results of the experiments, an angular criterion and a force criterion between the beam and the robotic bolting device were attained. They can be used for developing and improving the robotic assembly system.

2. Robotic bolting device

A robotic system used for construction automation is developed for bolt assembly automation[5]. In this study, we introduce a robotic bolting device that consists of a robotic manipulator and a bolting end-effector to assemble construction materials. For the robotic bolting device, a bolting mechanism using TS type bolt is employed, by which the bolting procedure can be fairly simplified.

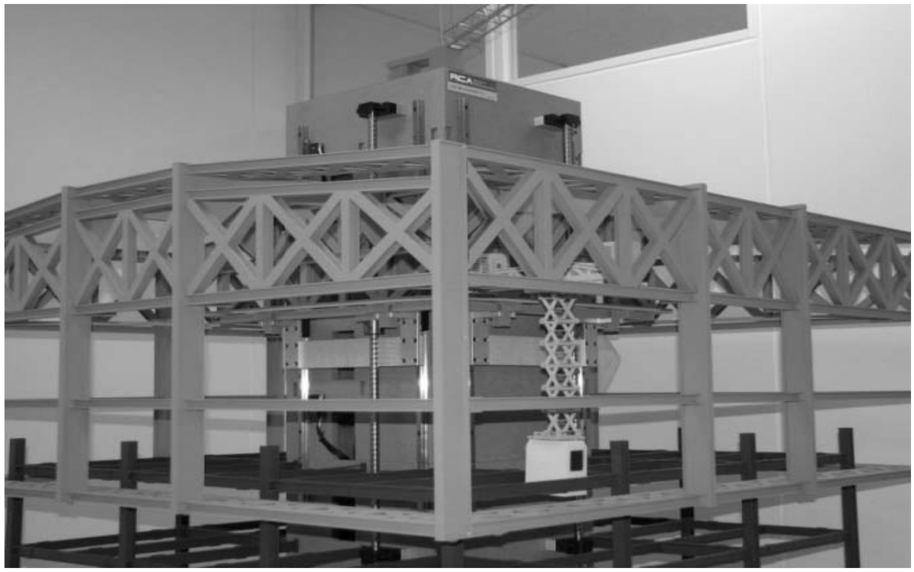


Figure 1. 1/20 size model of Construction Automation System of High-rise Building.

2.1 TS type bolt

Figure 2 shows a sample of TS type bolt which has a pintail and a breakneck at the end of the bolt. In a general bolting task, a worker should hold a hexagonal bolt head and rotate a nut against the bolt with a wrench. To use general bolts in the robotic bolting system, an additional device or more complex robotic systems are required. But by using TS type bolts in the robotic system, this problem is solved and lots of advantages are given to the robotic system. Since a bolting end-effector grabs both the bolt pintail and the nut from the pintail side at the same time and twist them to opposite directions, bolting operation is performed only at the pintail side. It simplifies the robotic system. Moreover, the pintail falls when bolt assembly ends, we can recognize the completion of the bolt assembly without other sensing devices.

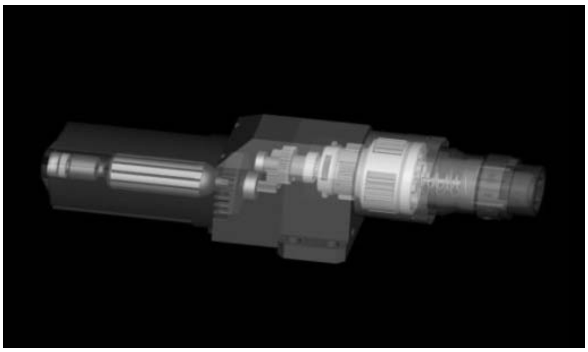
2.2 Bolting end-effector

A bolting end-effector is a device which operates the actual bolting task for TS type bolt. The bolting end-effector consists of DC motor, gear box, spring mechanism, and bolting parts. The bolting end-effector is actuated with a motor and a gear box to generate torque enough to fully assemble a bolt and nut pair. The pintail breaks when enough torgue is applied to TS type bolt and bolt assembly is completed. The pintail which remains in the bolting end-effector is detached outside when the bolting end-effector is pulled back. Figure 3 shows the 3D model and the prototype of the bolting end-effector.

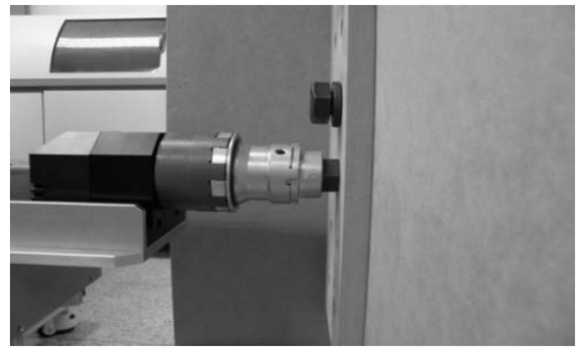
Figure 2. TS type bolt.

2.3 Robotic manipulator

The robotic manipulator has a gantry type robot mechanism which has 3-DOFs. The bolting end-effector is attached at the end of the z-axis frame. The bolting end-effector is moved to the region of interest using this robotic manipulator. Figure 4 is the prototype of the robotic manipulator. The three frames of the manipulation system fit to three dimensional Cartesian coordinates, so that it can be easily controlled with a simple control algorithm. Using this robotic manipulator, the workspace of 830*830*300 mm^2 can be guaranteed and the accuracy under dozens of micrometers can be achieved. With consideration of the clearance between the diameter of the bolt and the size of the hole being approximately 2 mm, the accuracy of the robotic manipulation system is satisfactory [6].



(a) 3D model of the bolting end-effector.



(b) Prototype of the bolting end-effector.

Figure 3. Bolting end-effector.

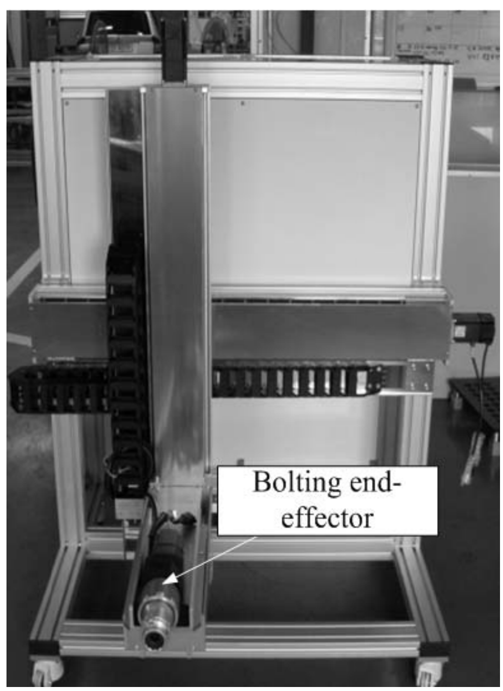
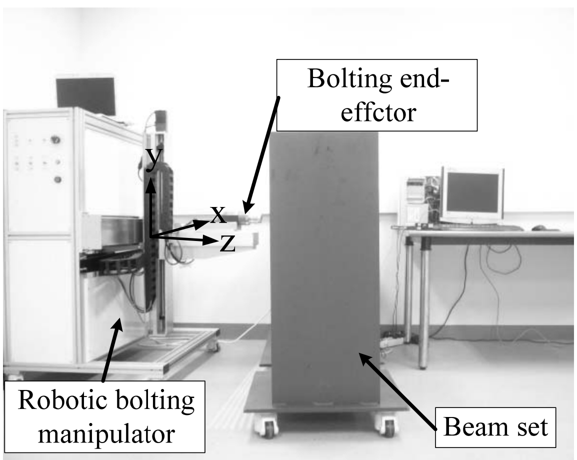


Figure 4. Prototype of the bolting manipulation system.

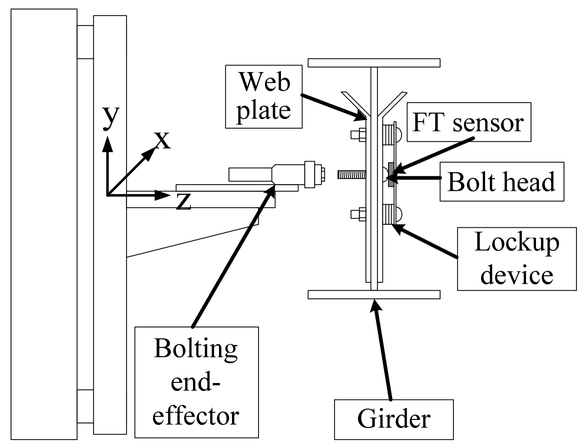
3. Experimental set for the robotic bolting assembly

Main purpose of this experiment is to determine the maximum feasible bolting angle between the girder and the robotic manipulator, and to examine supporting force of lockup device that prevents the bolt falling. Through the results of this experiment, we confirm feasible bolting criterions in bolting procedure using the robotic bolting device.

In order to replicate actual construction site in CF, an experimental set was made with a girder and web plate. The web plate is specially designed to align the beam and column automatically without an additional operation of an operator. Figure 5 (a) shows the experimental set which consists of the robotic bolting device, a girder/column assembly, and a FT(Force Torque) sensor. The FT sensor is attached to a lockup device that is placed on the opposite side of the robotic bolting device as shown in Figure 5 (b). The Bolting end-effector is placed on the bolting position by driving the x-y axis of the robotic bolting manipulator. And bolting is performed by proceeding z-axis of the gantry type robotic bolting manipulator and driving the bolting end-effector. We examined an acceptable bolting range between the girder and the robotic manipulator for successful bolting tasks. The FT sensor has direct contact with TS bolt head in order to measure the reaction force during the bolting process. Figure 6 shows relative angles between the girder and the robotic manipulator, where we marked 1 to 5 degree lines on the ground. Using these marks, we measured allowable relative angles between the girder and the robotic manipulator.



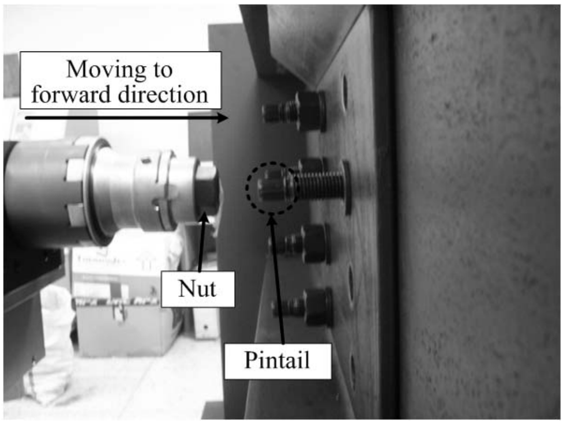
(a) Experimental set.



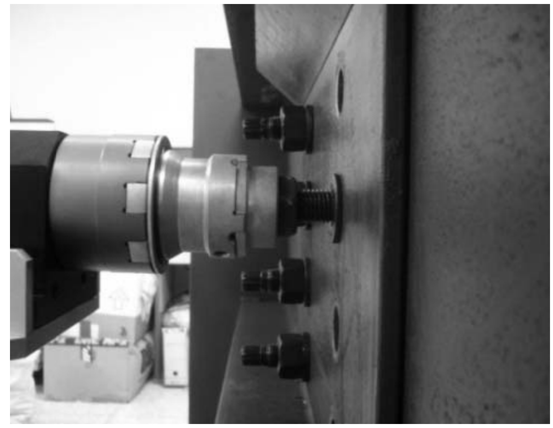
(b) Schematic diagram of experimental set.

Figure 5. Experimental set for measuring bolting reaction force.

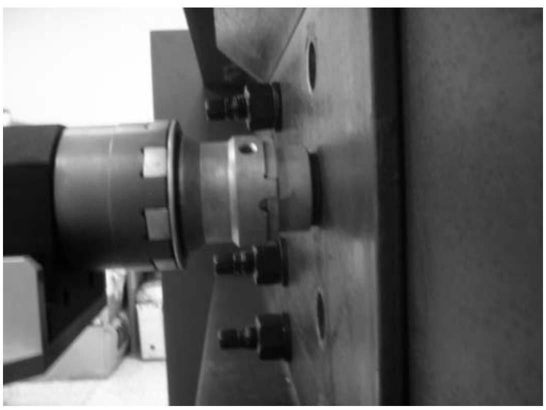
Figure 6. Relative angles between the girder and the robotic manipulator.



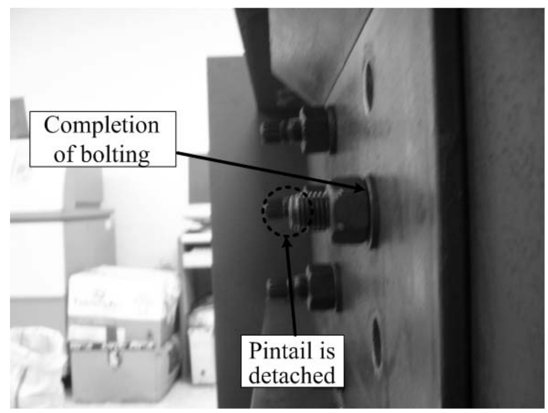
(a) The bolting end-effector is approaching.



(b) The pintail of the bolt is being inserted.



(c) the bolt and nut are being engaged.



(d) Bolting is completed.

Figure 7. Bolting assembly procedure.

4. Experiment and result

Figure 7 shows bolting experiment procedure. (a) The bolting end-effector is approaching the bolt. A nut is attached at the end of the bolting end-effector by magnet. (b) The pintail of the bolt is being inserted to the nut set on the bolting end-effector by proceeding z-axis of the gantry type robotic manipulator. (c) The bolting end-effector holds the nut and the pintail of the bolt at the same time, and twists them to opposite directions. And threads of the bolt and nut are engaged. (d) Bolting is completed and the pintail is detached. The proceeding speed of z-axis of the robotic manipulator is controlled to be the same as the feed rate of the bolt. Bolting experiments were repeated 10 times each with varying relative angles between the girder and the robotic manipulator from 0 degree to 5 degree.

Figure 8 shows reaction force measured by FT sensor with the relative angle of 4 degree. During Δt_1 ,, the bolting end-effector approaches the pintail of the bolt. Δt_2 is the period where the inserting and engaging task of the bolt and nut is performed. While the bolt and nut are being inserted and engaged, the z-directional reaction force increases and fluctuates. The force profile of Δt_3 in figure 8 shows that the reaction force is decreased after the engagement is completed. It means the bolting process is being done smoothly. In 4 degree of the relative angle between the girder and the bolting end-effector, the maximum reaction force to maintain the bolting task is 197N. Figure 9 (a) to (d) shows the bolting reaction forces according to the relative angles from 0 to 3 degree. In each case, the maximum reaction force is 48N, 101N, 124N, 155N. Reaction forces increase as the relative angle between the robotic manipulator and the girder increase. When the relative angle is more than 5 degree, the bolting task couldn't be performed. Therefore, the critical criterions for the successful bolting task using the robotic bolting manipulator are as follows: (1) The relative angle between the girder and the robotic manipulator is less than 5 degree. (2) The minimum force supporting the lockup device is more than 200N.

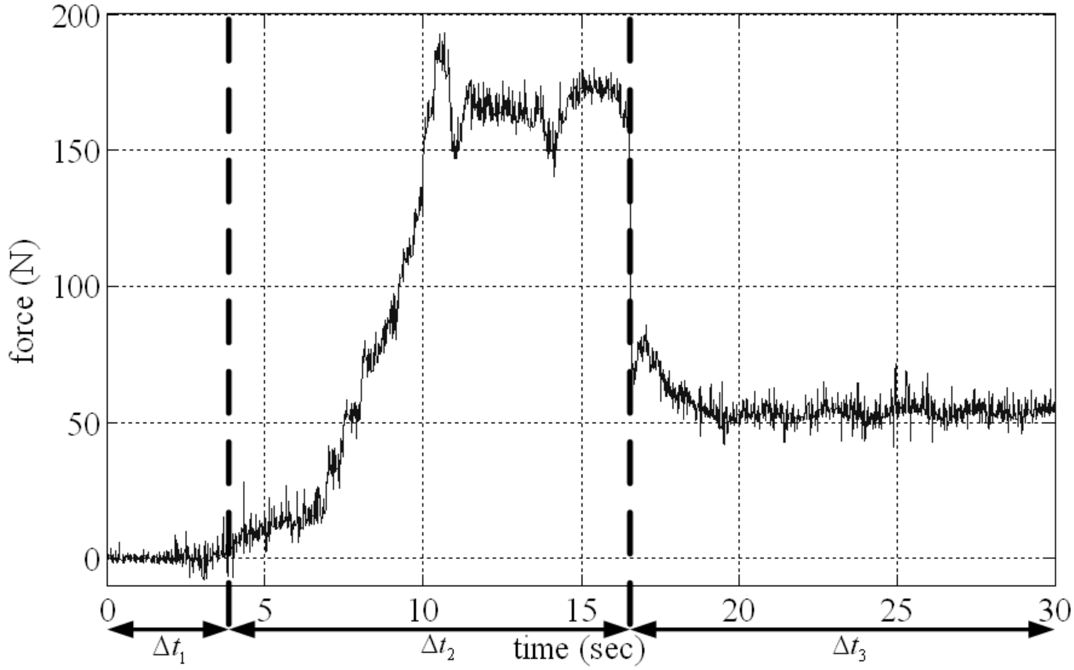
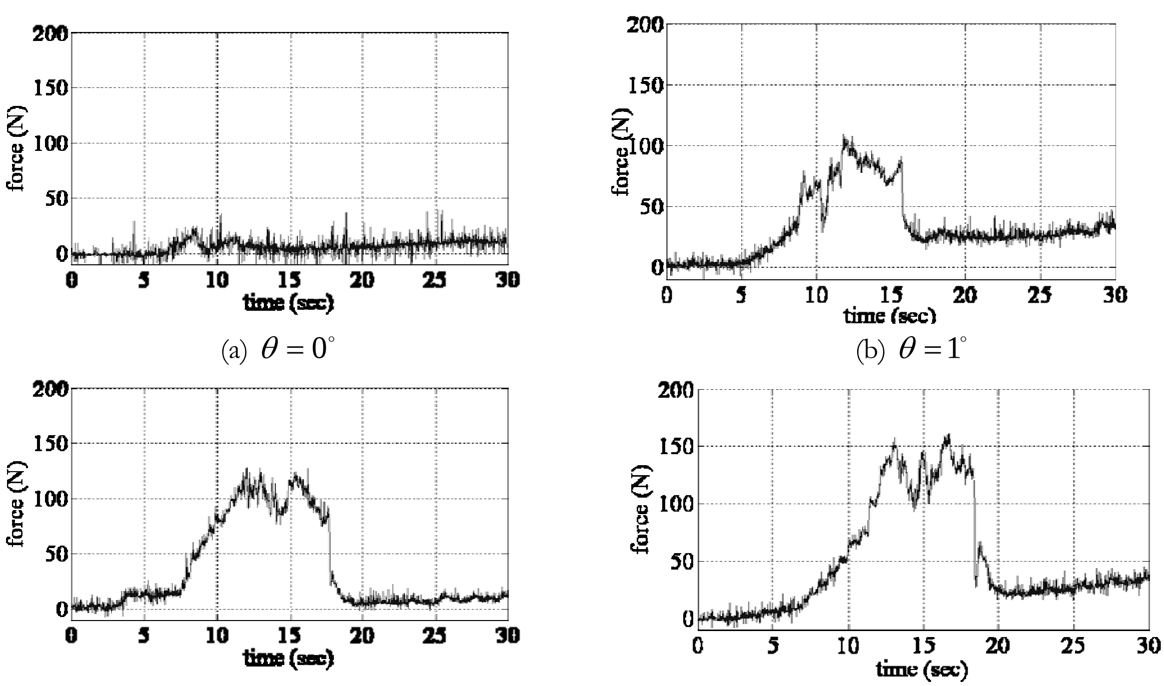


Figure 8. Experimental result. ($\theta = 4^{\circ}$)



(a) $\theta = 0^{\circ}$ (b) $\theta = 1^{\circ}$

(c) $\theta = 2^{\circ}$ (d) $\theta = 3^{\circ}$

Figure 9. Experimental results which has angles between the girder and robotic manipulator.

5. Conclusions

In this study, we obtained a force criterion and an angular criterion between the robotic bolting device and the beam to guarantee the successful bolting task. With various forces and angles, extensive experiments were performed to evaluate the feasibility of the bolting tasks using the suggested robotic bolting device.

The results attained by this research can be used to improve the performance of the developed robotic bolting device and develop an optimal version of the robot. Moreover, with employing the suggested bolting criterions, it is confirmed to control and operate the robotic bolting device safely and efficiently.

6. Acknowledgment

This research was supported by a grant (code# 06-D01) from Unified and Advanced Construction Technology Program funded by Ministry of Construction & Transportation of Korean Government.

References

[1] S. -K. Lee, N. L. Doh, G. -T. Park, K. -I. Kang, M. -T. Lim, D. Hong, S. Park, U. -K. Lee, and T. Kang, "Robotic technologies for the automation assemble of massive beams in high-rise building," Proceedings of International Conference on Control Automation and Systems, pp. 1209-1212, 2007.
[2] K. Jung, B. Chu, K. Bae, Y. Lee, D. Hong, S. Park, and M. -T. Lim, "Development of automation system for steel sonstruction based on robotic crane," International Conference on Smart Manufacturing Application, 2008.
[3] K. Hamada, N. Furuya, Y. Inoue, and T. Wakisaka, "Development of automated construction system for high-rise reinforced concrete buildings," Proceedings of IEEE International Conference on Robotics and Automation, vol. 3, no. 3, pp. 2428-2433, 1998.
[4] Y. Ikeda, and T. Harada, "Application of the automated building construction system using the conventional construction method together," Proceedings of 23rd International Symposium on Automation and Robotics in Construction, pp. 722-727, 2006.
[5] B. Chu, K. Jung, Y. Chu, D. Hong, M. –T. Lim, S. Park, Y. Lee, S. Lee, M. –C. Kim, and K. –H. Ko, "Robotic automation system for steel beam assembly in building construction," International Conference on Autonomous Robots and Agents, 2009.
[6] H. Nam, W. Choi, D. Ryu, Y. Lee, S. Lee, and B. Ryu, "Design of a bolting robot for constructing steel structure," International Conference on Control, Automation and Systems, pp.1945-1949, 2007.

FPGA-Based Real-Time Color Tracking for Robotic Formation Control

Ying-Hao Yu[1], N. M. Kwok[2] and Q. P. Ha[1]

[1]School of Electrical, Mechanical and Mechatronic Systems, University of Technology, Sydney, Broadway, NSW 2007, Australia, Email: {yinghao.yu, quangha}@eng.uts.edu.au
[2]School of Mechanical and Manufacturing Engineering, University of New South Wales, Sydney, NSW 2052, Australia, Email: nmkwok@unsw.edu.au

Abstract

In construction automation, the tracking of worker location or a moving object is important for labor monitoring, resource management, and machine coordination. For object tracking, a camera is often utilized to obtain information of construction vehicle motion which can then be employed in coordination controls. Real-time tracking of autonomous vehicles, particularly for the control of multiple targets in formations, still suffers from constraints imposed in computation resources. Here the field programmable gate array (FPGA) technology is applied in a prototypical tracking system for vehicles by using a CMOS camera to detect their color-tags. The raw image from the Bayer color pattern is used to indicate the 2-dimensional position of vehicles and encrypted infrared commands are issued to deploy them in a leader-follower formation. It is shown that the novel system-on-programmable-chip with parallel control cores design can efficiently handle color recognition and multi-vehicle control while significantly reduces memory requirement and computation time.

Introduction

The digital camera has become a popular sensor for monitoring and surveillance systems. By improvement of the semiconductor technology, modern digital cameras with high pixel number can provide greater image detail in various applications. A digital image sensor usually utilizes an array consists of the charged-couple device (CCD) or complementary metal oxide semiconductor (CMOS). By passing light through the lens and color filter array (CFA), the real image is transformed into the mosaic-like RGB image projected on the digital image sensor array (Lukac and Plataniotis, 2007). The raw RGB image then is reconstructed as a meaningful image for human perception by de-mosaicking (Lee, 2005) and color correction (Gonzalez and Woods, 2002). For sensing applications, the additional image processing to extract object's color strength or contour (Nixon and Aguado, 2008) is required for every image pixel. Moreover, it should be noted that imaging processing consumes a lot of memory space and most of computation resources in a computer system.

The digital camera is also a popular sensor in a multi-robot system (Fierro et al., 2002) for obstacle detection and environment learning. Such a system is a very valuable tool in the development of an automated coordination system for construction vehicles. In a real-time object tracking and detection system, particularly for moving targets, the picture acquisition and image processing speed critically determines the system performance. Low computation speed will cause the loss of precise driving control for the multi-robot system and results in collisions. When the robot system needs a high resolution camera to improve imaging details of the object, the coherent time delay with massive computation for image processing will occur. On the other hand, even the powerful microprocessor can mitigate time delay, the higher power consumption is also unsuitable for a mini robot system with battery powered operation (Bräunl, 2006).

Thus a compromise scheme for the multi-vehicle system in image processing issues is sharing the image processing and strategy making tasks in an external server (Fierro et al., 2002). Therefore, the vehicle carries only a simple embedded system which captures and sends image data to the server by wireless in order to get a balance between power consumption and execution speed. Based on the same idea, we propose a new colour-based tracking system for vehicle coordination using color-tag recognition for identification with an external server. This scheme will achieve higher performance of power management and real-time control speed.

In our design, each vehicle will have an embedded system for motion control and communicate by the external server by an encrypted infrared signal. Here, to avoid the access to a large external storage device when running the whole program and acquiring information from outside for the control purpose, we propose use the Field Programmable Gate Array (FPGA) technology (Yu et al., 2008). The server is designed by using an FPGA chip with parallel logic control groups (cores) in order to achieve such benefits of low power, low cost, and flexible speed control as compared to a general purpose computer. The designed external server will monitor the vehicle by the captured 2-dimentional object positions with a single digital camera equipped on the server. The target tracking by using raw image, window of interest (Carvalho et al., 2000), and the dynamic trust region (Wang et al., 2006) designs is obtained for a low computation burden in the FPGA chip. Finally, the low logic gate usage and the multi-object tracking with the basic leader-following strategy in formation control will demonstrate the feasibility of this novel system.

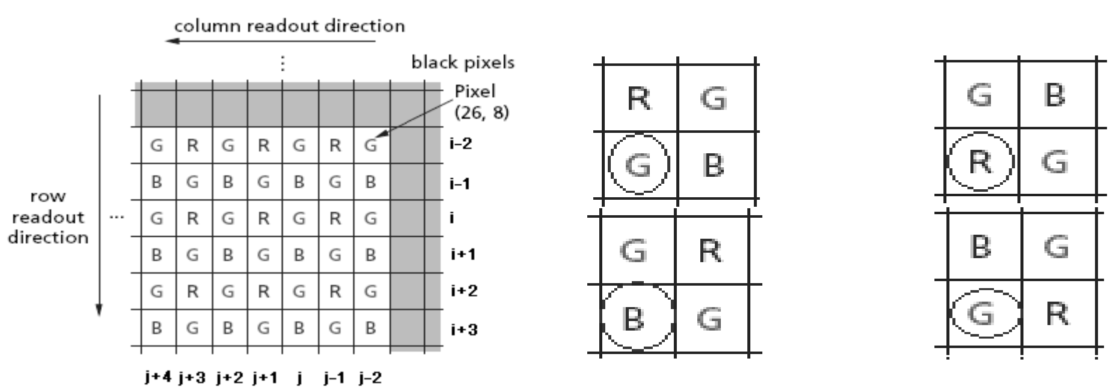
The paper is organized as follows. Section II presents the raw image processing with Bayer pattern, color prediction, and noise filtering. The color tracking mechanism is described in Section III. The FPGA design, resource usage, and the experimental result captured from the video images are illustrated in Section IV. Some discussion of the development is given in Section V. Finally, a conclusion is drawn in section VI.

Processing of Raw Image

Image processing with large matrix calculation will deteriorate the system performance due to limitations in computation speed. The simpler image processing such as the threshold (Hu et al., 2006) could be the better choice for lower demand for computation resource. Based on this strategy, we develop the color prediction and noise filter algorithm with the raw image to achieve real-time color recognition.

Bayer Pattern

The mosaic-like RGB image in the Bayer pattern (Lukac et al., 2005) is shown in Figure 1.



(a). The Bayer pattern (b). Possible combinations for every four pixels

Figure 1. Bayer pattern of a typical digital image.

The color pixels are obtained by scanning the image line by line. We determine the color components of the circled pixels, Figure 1(b), by a local neighborhood of every four scanned pixels. In our system, we only need to recognize whether the label of object's surface belongs to a blue or green color range. Hence, we define the basic threshold range for blue- and green-like color as:

$$G \Leftrightarrow (g_1 \cap g_2) > (r \cup b) \quad (1)$$

$$B \Leftrightarrow b > (r \cup g_1 \cup g_2), \quad (2)$$

where g, r, and b denote the RGB pixel strength for the combination of four color pixel units.

Color Prediction

In real world problems, the object's color always changes due to effects of light source, shadow, and reflection on surface. The basic color definition above is, therefore, not sufficient for an arbitrary environment. Thus, the color prediction is designed to compensate for the incurred color error. We rewrite the color definition as:

$$G \Leftrightarrow (g_1 \cap g_2) > (r \pm \Delta r_n \cup b \pm \Delta b_n), \text{ if } G \Leftrightarrow g_{1(i-1,j)} \cap g_{2(i-1,j)} \quad (3)$$

$$B \Leftrightarrow b > (r \pm \Delta r_n \cup (g_1 \cup g_2) \pm \Delta g_n), \text{ if } B \Leftrightarrow b_{(i-1,j)}, \quad (4)$$

where Δr_n, Δb_n, and Δg_n are the maximum color strength errors that are measured from the object's surface in the test environment by the digital sensor, n denotes the different color strength level below the saturation value. The comparison between color prediction and image captured in our laboratory is given in Figure 2. The system shows the full scale of grey if the green color is confirmed. Figure 2(b) shows the grey levels of detecting green where uncertainties exhibit due to the shadow in image 2(a). In the contrary, Figure 2(c) illustrates a good detection result with color prediction by using (3).

(a) (b) (c)

Figure 2. The operation and comparison of the color prediction approach.

Noise Filter

Since color prediction can make the full grey level image more complete, the background noise can be eliminated by examining the continuity of pixel groups. We add an extra criterion defined by:

$$G \Leftrightarrow G_{l1(i,j)} \cap G_{l2(i,j)} \cap G_{l3(i,j)} \quad (5)$$

$$B \Leftrightarrow B_{l1(i,j)} \cap B_{l2(i,j)} \cap B_{l3(i,j)}, \quad (6)$$

where Gl(i,j) are the active pixel on scan line "l". Figure 3 shows the performance of noise filtering combined with the color prediction function. The mini robot's grey image is improved for color tracking tasks with the proposed color prediction and noise filter as shown in Figure 3(c). Here, a mini-robot, the Eyebot (Bräunl, 2006), was used in laboratorial experiments to test the development for color detection purposes on a construction site.

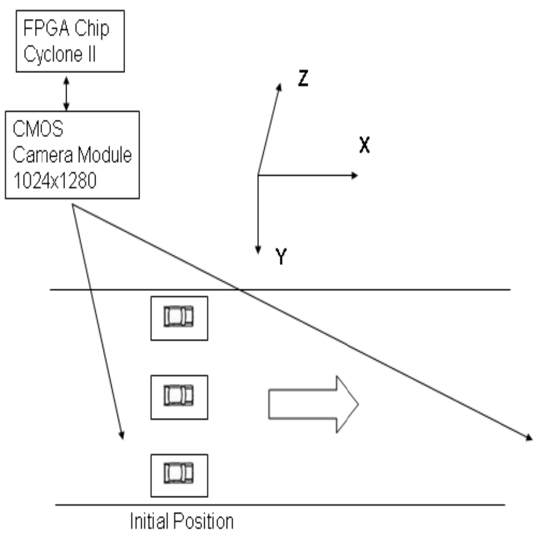
(a) Eyebot (b) Robot image surrounded by noise (c) Image with noise filter.

Figure 3. The performance of noise filtering.

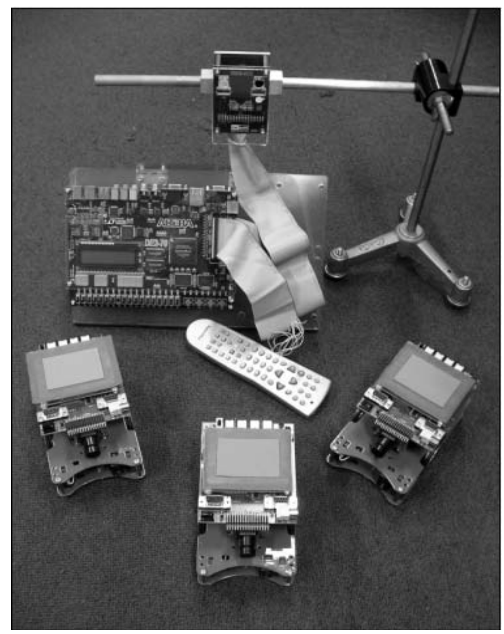
Color Tracking Mechanism

The color tracking system is controlled by an external server which is designed on a FPGA development platform, the DE2-70 with Cyclone II FPGA, from Altera®. Figure 4 shows the scheme and the real system components. The FPGA chip monitors the mini robot, via a single camera above the robots. For activating the tracking mechanism, every Eyebot needs firstly to pass through the initial area and then monitored by the FPGA chip. The chip will recognize the robot via a 6 x 4 cm2 square greenish blue paper adhered on top as the window of interest (Carvalho et al., 2000); the robot will be detected and tracked if the counted blue pixel amount is over the threshold level in the initial docking area.

The green cloth attached with the greenish blue paper is used to initialize the start and end point of pixel counting. After the Eyebot is locked by the FPGA chip, the chip can track the object by defining the dynamic trust region. In every dynamic trust region, the FPGA chip checks the threshold level of sampled blue pixel amount and drawing the new red trust region by the moving centre point in blue pixel area. Here we assume the maximum speed of centre point is moving for 1/2 length on x or y coordinate of last tracked blue area in rectangular shape then set the new trust region for four times bigger than the blue label as the minimum tracking dimension with the Eyebot's speed 3.8 km/ hr. Finally, the FPGA platform contacts the Eyebots via the encrypted infrared commands of a TV remote.



(a) Color tracking scheme



(b) Components of the vehicle/robot tracking system

Figure 4. Color tracking system and components.

FPGA Chip Design and Experimental Scenario

Unlike the general purpose processor which adopts fixed hardware design for the maximum compatibility and flexibility, we embed the logical programming into the hardware circuit design directly. This kind of design helps the low speed system to satisfy the demand for sufficient real-time control with highly parallel architecture, shown in Figure 5. The raw image is distributed to three control groups (cores), each group controls one Eyebot and the leader core endowed with the extra ability to lead the followers.

An Altera® FPGA platform supplied with Terasic® auxiliary VGA demo program is used in our experiments. The surveillance camera is installed up from the ground for 1.2 m with 45 degrees inclined to the Eyebot on ground, and the frame rate is 12 fps for 1024×1280 resolution. The monitoring area under stationary camera monitoring is 1.2×1.6 m2.

Figure 6 presents the test results captured from the camera. The test scenario is shown in Figure 6(a). The FPGA chip locks three Eyebots immediately after the robots enter initial docking areas, Figure 6(b) to 6(c). In our test scenario, the leader drives forward automatically after a period of time delay, then the two followers move forward immediately after they received the commands from the leader, Figure 6(d). In the end, the leader waits for the followers at the goal area and asks them to stop after having reached the goal area, Figure 6(e). This test scenario demonstrates the use of color tracking, remote wireless polling, parallel driving control, and basic multi-robot formation (a wedge), performed by hardware design in a single chip.

The chip resource usage given is summarized in Figure 7. It is shown that the novel color tracking system including the VGA demo interface only consumes 7% of the total logic elements (LEs). This saving in resources would enable us to put in more intelligent control functionalities in the single chip. This advantage is rather important in the robotic formation control in the presence of obstacles and also for construction automation where the tracking of worker location or a moving object is important for labor monitoring, resource management, or machine coordination.

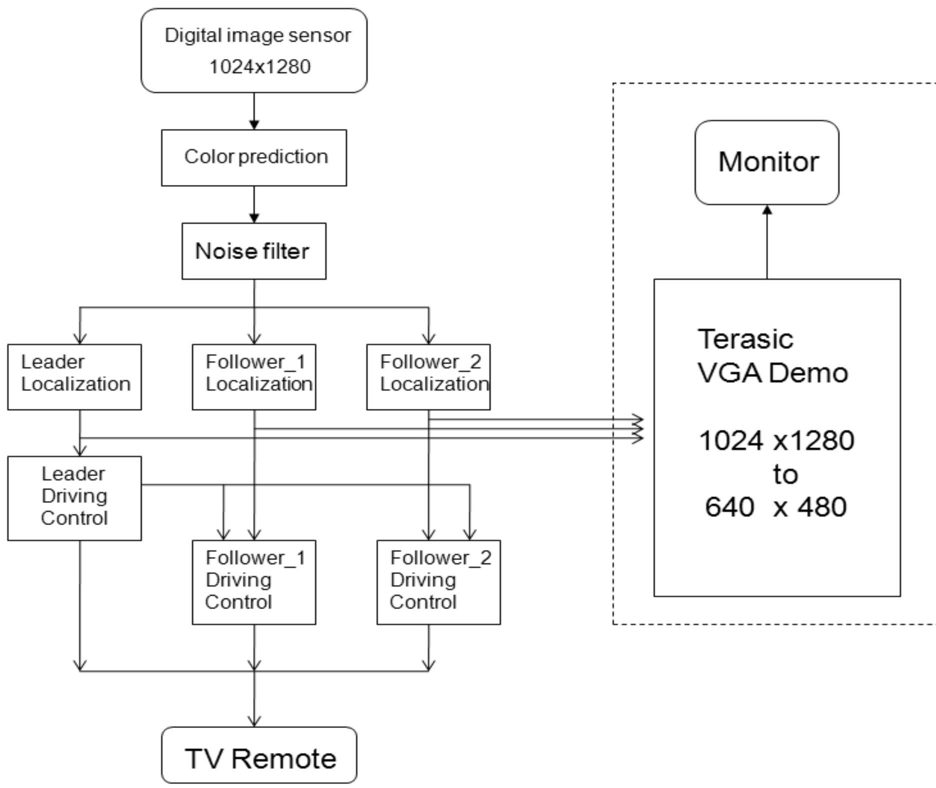


Figure 5. Internal function design of the FPGA chip.

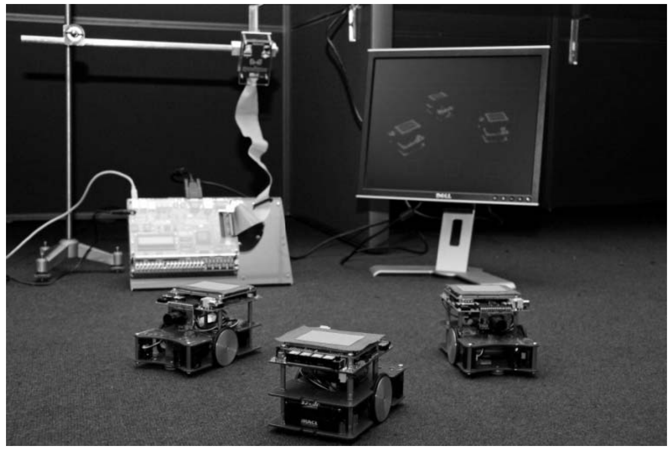
Discussion

The proposed color tracking approach using raw image, initial docking position, and dynamic trust area has contributed to a good performance in the proposed color tracking system. The color discrimination is useful in detecting the worker's vest, moving vehicle, and the position of conveyer for operation or security reasons, which is essential for construction automation.

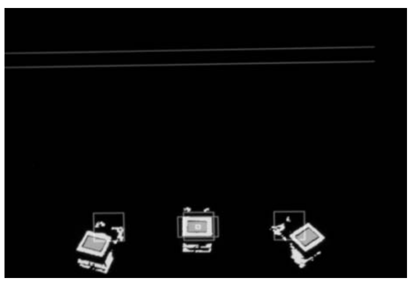
Meanwhile, the color prediction plays a critical role by providing a reliable image for object recognition, and the real-time raw image processing is not effected by the complexity of high resolution digital sensor array. Notwithstanding the automatic facilities usually need to operate in the adverse circumstance in construction site, the system-on-programmable-chip design with FPGA chip can provide the superiority for easy installation, maintenance, and stability. The savings in chip resources allow more room for further incorporation of decision making and control algorithms.

Conclusions

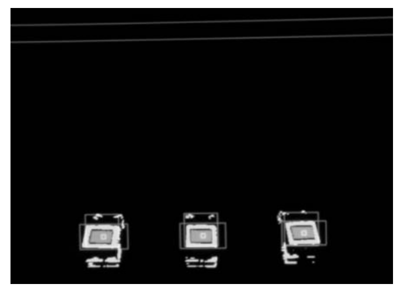
This paper has presented a real-time color tracking system with low cost and low power consumption. The refined circuit design with raw image, initial docking detection, dynamic trust area tracking, and parallel control guarantee satisfactory performance in FPGA chip design. The effectiveness of the approach has been demonstrated by experiments conducted using mini-robots. Finally, the least LE usage in FPGA chip also inspires more intelligent control design in the future. The development is useful in the tracking of worker location or a moving object labor monitoring, resource management, or machine coordination for construction automation.



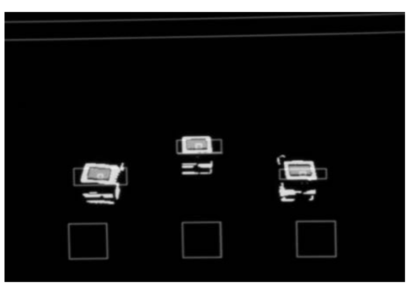
(a) Test scenario



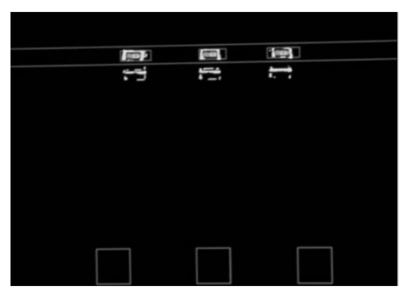
(b) Eyebots entering initial docking areas.



(c) Eyebots locked in docking areas.



(d) Wedge formation: one leader, two followers.



(e) Leader and followers reaching the goal area.

Figure 6. Color tracking with three Eyebots.

Flow Status	Successful - Wed Dec 10 20:32:59 2008
Quartus II Version	7.2 Build 151 09/26/2007 SJ Full Version
Revision Name	
Top-level Entity Name	
Family	Cyclone II
Device	EP2C70F896C6
Timing Models	Final
Met timing requirements	Yes
Total logic elements	4,595 / 68,416 (7 %)
Total combinational functions	4,392 / 68,416 (6 %)
Dedicated logic registers	1,941 / 68,416 (3 %)
Total registers	1941
Total pins	218 / 622 (35 %)
Total virtual pins	0
Total memory bits	82,852 / 1,152,000 (7 %)
Embedded Multiplier 9-bit elements	0 / 300 (0 %)
Total PLLs	1 / 4 (25 %)

Figure 7. Resource usage in the FPGA chip.

Acknowledgement

This work is supported by UTS research grant 2006000848 and, in part, by the Centre of Excellence programme, funded by the Australian Research Council (ARC) and the New South Wales State Government.

References

[1] Bräunl T. (2006). Embedded Robotics, Springer-Verlag New York, Secancus Inc., NJ, USA.
[2] Carvalho José R.H., Rives P., Ailton S.B, and Bueno Samuel S. (2000). "Visual Servo Control for a Class of Mobile Robot," Proc. 2000 IEEE Intl. Conf. on Control Application, Anchorage, AK, USA, pp. 431-436.
[3] Fierro, R., Das, A., Spetzer, J., Esposito, J., Kumar, V. , Ostrowski, J., Pappas, G., Taylor, C., Hur,Y., Alur, R., Lee, I., Grudic, G. and Southall, B. (2002). "A Framework and Architecture for Multi-Robot Coordination," International Journal of Robotics Research, Vol.21, No. 10-11, pp. 977-995.
[4] Gonzalez R.C. and Woods R.E. (2002). Digital Image Processing, Prentice Hall, Upper Saddle River, NJ, USA.
[5] Hu Q., Hou Z., and Nowinski Wieslaw L. (2006). "Supervised Range-Constrained Thresholding," IEEE Trans. on Image Processing, Vol. 15, No.1, pp. 228-240.
[6] Lee X. (2005). "Demosaicing by Successive Approximation," IEEE Trans. on Image Processing, Vol. 14, No. 3, pp. 370-379.
[7] Lukac R. and Plataniotis K.N. (2007). Color Image Processing: Methods and Applications, CRC Press.
[8] Lukac R., Plataniotis K.N. and Hatzinakos D. (2005). "Color Image Zooming on the Bayer Pattern," IEEE Trans. on Circuits and Systems for Video Technology, Vol. 15, No. 11, pp. 1475-1492.
[9] Nixon M.S. and Aguado A.S. (2008). Feature Extraction and Image Processing, Academic Press.
[10] Wang Z., Zhang K., Wu Y., and Lu H. (2006). "Tracking Dim Target in Infrared Imagery Using the Trust Region Embedded Particle Filter." Proc. 2006 8th Intl. Conf. on Signal Processing, doi 10.1109/ICOSP.2006.345667.
[11] Yu Y.-H., Kwok N.M. and Ha Q.P. (2008)., "A New System-On-Programmable-Chip (SOPC) for Smart Car Braking Control," Proc. Int. Conference on Intelligent Technologies (InTech 2008), Samui, Thailand, 7-10 Oct., pp. 34-39.

Computer-aided Methods for Operations Planning

Prof. Dr.-Ing. Peter D.P. Böttcher

Department of Construction Management, Hochschule für Technik und Wirtschaft des Saarlandes, Goebenstr. 40, D-66117 Saarbrücken, phone +49-681-5867230, boettcher@htw-saarland.de, www.htw-saarland.de/fb-bi/fg/baubetrieb

Abstract

In order to enable the use of semi-automatic and automatic operations for production on building sites, operations must be planned in advance and controlled on the basis of the planning.

In a current research project funded by the Federal Republic of Germany, a method is being developed, which supports the computer-aided planning and control of medium-sized building sites. Works on building sites of this size take 4 to 12 months, and they are mostly planned insufficiently. This insufficient planning is due to a lacking support of construction management in the preparation of the building site.

The method is aimed at site managers who manage several medium-sized building sections within one large construction site. Planning of operations and building site facilities is supported as well as economic monitoring via hour and cost control. The required actual data is acquired electronically by the building site staff.

The method incorporates the usual workflow of the site manager and combines the information into few essential characteristic figures.

Keywords: knowledge management, planning of construction projects, economic, workflow

1. Introduction

The German building industry is strongly characterised by medium-sized businesses. Of approx. 76,000 companies in Germany, about 7,700 have between 20 and 200 employees. These companies account for approx. 52% of the € 78 billion annual turnover in the building industry and thus have a very high economic significance in the German construction sector.

On the other hand, companies of this size are often neglected when it comes to improving productivity, because they do not provide enough economic potential to develop their own programmes. Thus, the responsible site managers often still work with old methods for planning and controlling building sites and coordinate the sites insufficiently.

The current research project wants to develop methods in this area which support the working methods of site managers in order to achieve improved profitability.

The analysis of the workflows in companies of this size showed the following deficits:

a. There are no systematic methods for planning and controlling building sites within the organisation. The workflows are often designed too much for building companies with more than 3,000 employees.
b. There are no tools for persons who manage three to five building sites at the same time and generally do not have a group of employees for managing the building site. It is true here, too, that the majority of tools are designed for large construction sites with an individual group of employees for managing the site.

2. Requirements of the Organisation

Building companies with 20 to 200 employees have widely differing structures for the organisation of workflows regarding order processing and for the planning and controlling of the building site. Generally, the organisational structure developed over the years and only rarely was changed systematically but rather adjusted to the day-to-day requirements.

Structures have developed which generally meet the requirements of every-day work. The companies do see room for improvement; however, they only want improvements to be made if the organisational principle of the respective company does not have to be changed too much.

The workflows in planning and controlling building sites shall be examined. These include the "planning of the building site", the "acquisition of information on the building site" and the "comparison of the planning with the actual operations on the building site" [4],[7], [10].

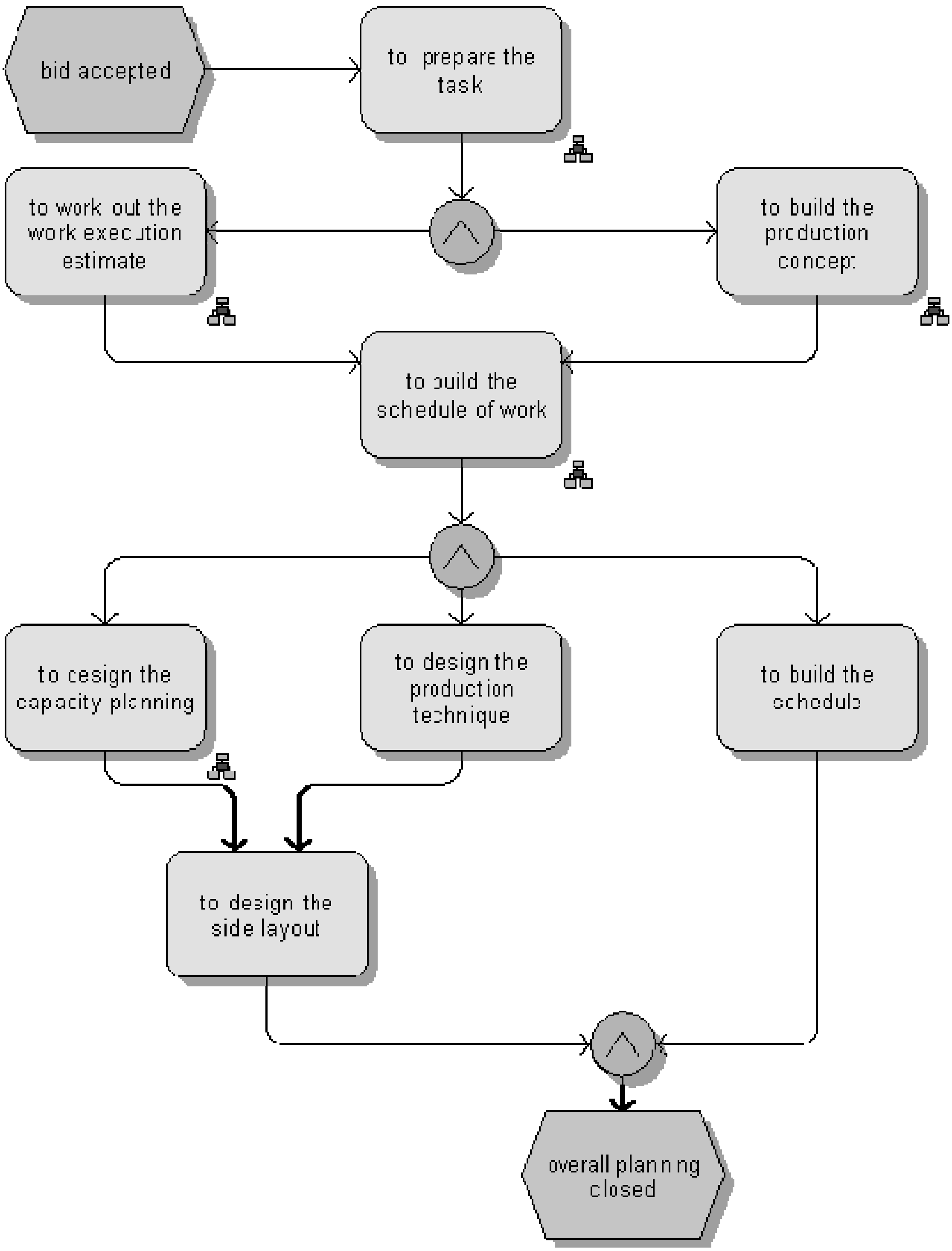


Fig. 1: Basic model "Set up operational planning" as event-driven process chain

3. Systematic Methods

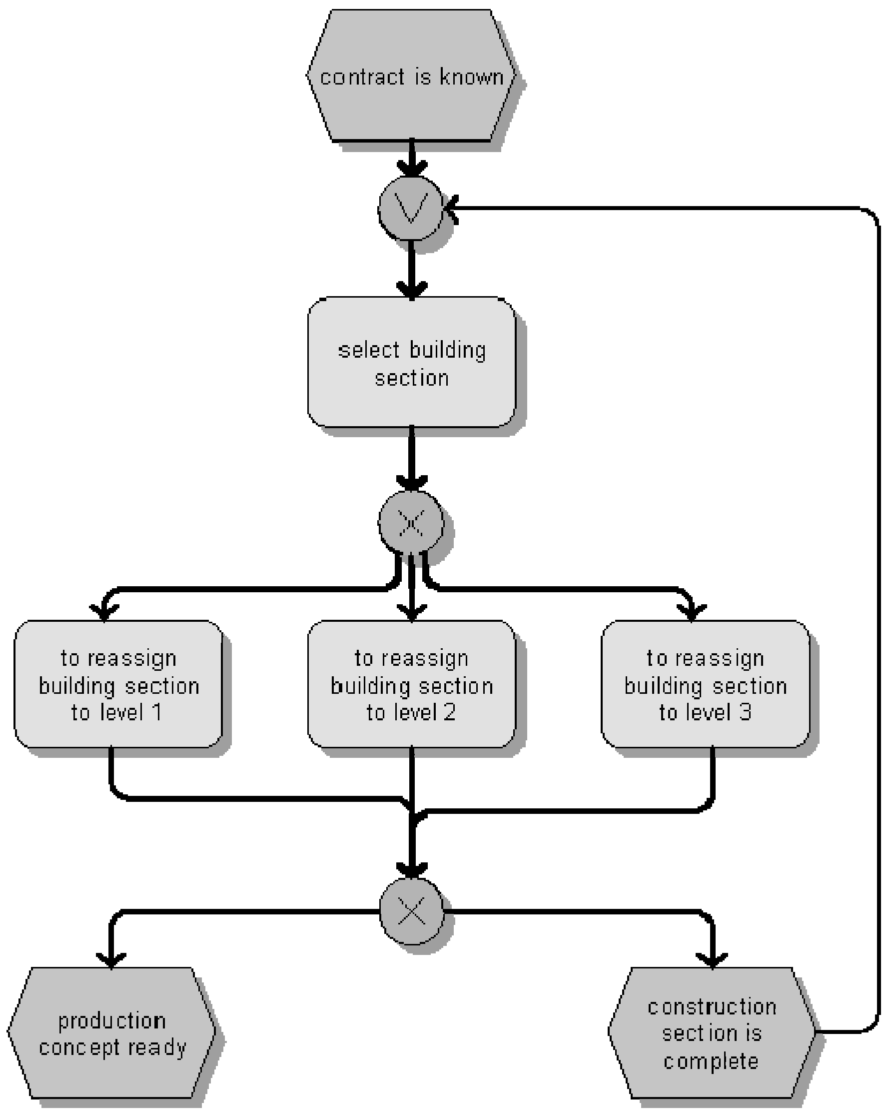


Fig. 2: Model "Set up production concept"

In order to be able to evaluate the various requirements of the companies, a reference model was established first (Fig. 1). This model is the basis, with which the task "Set up operational planning" (Fig. 2) is compared to the organisational structures of the companies. In this comparison, companies of various sizes and order structures are examined.

The reference model is modelled on the basis of "event-driven process chains" (EPC); the current development state of the model can be viewed at the Internet1.

Models on the basis of the EPC [6] consist of four components. The task (function) is symbolised by a green square, the work result (event) by a red hexagon, decisions by a grey circle with the decision characteristics (X) for either/or, (□) for several variants possible and (V) for all variants must be selected. The fourth element is represented by the directional edges, which specify the organisational order.

The individual functions in the basic model are refined by additional models. The refinements are regarded as options. Not all tasks in a refinement are implemented by the companies, only those that are desired.

[1] http://bi-baukom.htw-saarland.de:8080/businesspublisher

However, it is possible to examine the workflow in the company on the basis of the reference model, and weaknesses are shown if provided.

A strongly simplified guideline for smaller building companies has already been developed from the reference model.

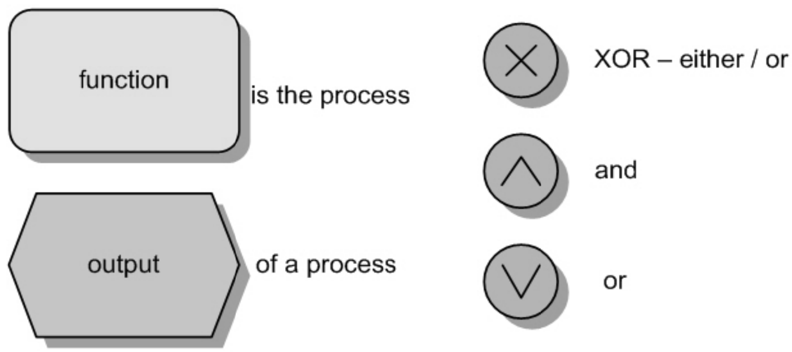


Fig. 3: Model "Set up production concept"

4. Computer-aided Tools

The operational planning of a building site contains the tasks "Planning and preparing the building site" and "Controlling the building site". Both task areas are covered by the site manager responsible and generally without support by additional staff. If it is intended to improve the processing of these task areas and increase the productivity of the building construction in a second step by means of automatic or semi-automatic production facilities, processing must be carried out with the aid of computers. The basis for the development of a computer-aided tool is the reference model, which is provided with detailed instructions and calculations.

The tasks "Generate production concept", "Generate work directory", "Acquire data on building site" and "Short-term income statements" have proven to be critical in the day-to-day implementation of the site manager. Particularly the structure according to which a site manager manages the building site varies a lot and has an impact on all four tasks. If the computer-aided tool is not flexible enough for the implementation of these four tasks, the tool is not used.

4.1 Generating a production concept

In the production concept, the site manager thinks about how to divide the building project into building sections in order to achieve an optimum workflow. The development of a production concept is a creative thought process and must be supported as such by the software and not hindered. The organisation of the building sections, the assignment of sub-sections must be easy to alter. Table calculation programs, for example, are awkward to handle; in contrast, the table structures in word processing programs are very easy to handle. [3]

4.2 Generating a work directory

The work directory collects the services that need to be performed in a building section. Furthermore, the planned revenues and the target hours for the building section are displayed. The work directory thus is an element of the target-performance comparison to be made at a later point.

However, the display of the planned revenues on the basis of the calculation also is the problem when generating a work directory. It is required that all services are recorded and assigned to the building sections. This work cannot be performed by the site manager before construction starts.

Therefore, a method was developed for the work directory with which the revenues per building section can be estimated although only the essential services were recorded. In the course of the building project, the site manager can add the missing services, thus receiving more accurate information on the building section. [5], [8], [2]

4.3 Acquiring data on the building site

In order to be able to compare the planned revenues with the actual costs, the data for the cost calculation

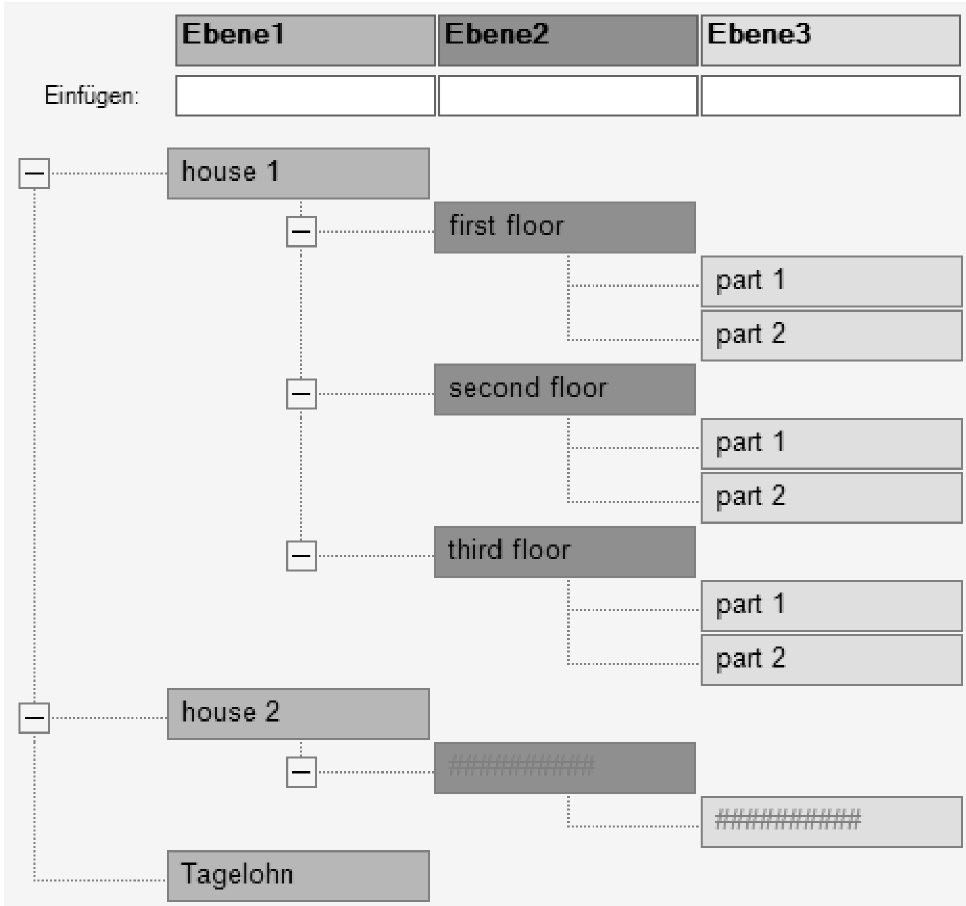


Fig. 4: Design tool production concept

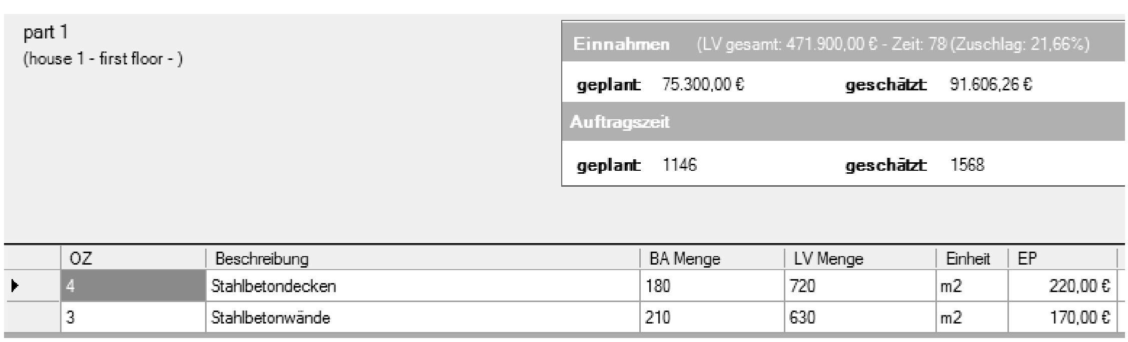


Fig. 5: Work directory

must be acquired on the building site. The data can only be acquired by the building site staff.

However, the quality of the data acquisition relies on the fact that the person acquiring the data knows what the data is being acquired for. This generally is not given on building sites. The foreman has to acquire data, but has nothing to do with their evaluation. Therefore the quality of the data is insufficient.

Thus a procedure was developed by means of which the foreman receives information on the economic situation of the building site. The data is acquired by means of handheld PDAs. The software on the PDA is set up in such a way that first the building section and then the staff working on the building section is recorded. As a result, the foreman only has to acquire the data when he assigns the employees to a new building section. At the same time, the foreman can see the current hours worked on the building section.

Fig. 6: Data acquisition on the PDA

4.4 *Short-term income statements*

The information on the building section is essential for the site manager for managing a construction site. Experiences can be transferred from one building section to another. The comparison of the planned revenues from the work directory with the actual data from the data acquisition has to be up-to-date and made individually for each building section.

For this, the data from the data acquisition has to be combined into characteristic figures first. The costs of the building site generally consist of costs for wages, material, production facilities and subcontractors. Since the paid working hours are recorded very accurately, the wage costs can be calculated easily; the material costs are the second large area of expenses. However, not all materials can be recorded and calculated in a timely manner. Here, it is necessary to make estimates on the basis of the most important materials. The subcontractor services can be recorded via checked invoices. The production facilities must be estimated by means of comparing the costs with already completed building sites.

building section:	house 1 first floor part 1						
labor cost		1350	h	25,00 €	/h		33.750,00 €
material cost							34.750,00 €
	concrete	120	m3	65,00 €		7.800,00 €	
	steel	80	t	250,00 €		20.000,00 €	
						27.800,00 €	
	loading	25,0%				6.950,00 €	
subcontractor cost							2.500,00 €
equipment cost		6,0%					4.260,00 €
overhead cost		20,0%					15.052,00 €
						expected cost:	90.312,00 €

Fig. 7: Cost structure for a building section

Studies have shown that the estimated costs per building section can achieve an accuracy of 95%. The advantage is the timely evaluation of the building section. Generally, the costs can be calculated one day after the building section has been completed, and the site manager can transfer the results to the next building section.

building section			set income	actual cost	margin
house 1			240.606,00 €	173.312,00 €	- 1.706,00 €
	frist floor		126.606,00 €	90.312,00 €	1.294,00 €
		part 1	91.606,00 €	90.312,00 €	1.294,00 €
		part 2	35.000,00 €		- €
	secound floor		114.000,00 €	83.000,00 €	- 3.000,00 €
		part 2	80.000,00 €	83.000,00 €	- 3.000,00 €
		part 2	34.000,00 €		- €
	third floor		- €	- €	- €
		part 1			- €
		part 2			- €

Fig. 8: Target-performance comparison of the building sections of a building site

In the target-performance comparison, two estimated values are compared on the basis of the finished building section. The revenues to be expected result from the work directory, and the probable costs result from the data acquisition with the corresponding calculation mode. The values are sufficient for evaluating a building section, because particularly the fast acquisition of the characteristic figures and the transfer to other building sections is important here.[3],[9]

References

[1] Böttcher, P. Santowski, G. Hitzel, A. and Gerster, R. (Hrsg.), (2007), „IFF bau´07 – Innovation und Forschung an Fachhochschulen „Baubetrieb und Bauwirtschaft“, Fortschritt-Berichte VDI Reihe 4 Nr. 207, VDI-Verlag Düsseldorf

[2] Böttcher, Peter (2003), „Arbeitszeitrichtwerte im Zimmerhandwerk“, Bruderverlag Karlsruhe

[3] Böttcher, Peter (2007), „Produktionsmanagement in klein- und mittelständischen Baufirmen“, Tagungsband: Innovation und Forschung an Fachhochschulen, Fortschritt-Berichte VDI Reihe 4 Nr. 207, VDI-Verlag Düsseldorf

[4] Feledziak, Jerome (2005), „Rechnergestützte Werkzeuge im schlüsselfertigen Bauen“, Deutsch-Französische Hochschulinstitut der HTW des Saarlandes, Diplomarbeit

[5] Gerster and Kohl (2003), „Baubetrieb in Beispielen“, Werner-Verlag, Düsseldorf

[6] Keller, Nüttgens and Scheer: Semantische Prozeßmodellierung auf der Grundlage Ereignisgesteuerter Prozeßketten (EPK). Erschienen in der Reihe: Veröffentlichungen des Instituts für Wirtschaftsinformatik. A.-W. Scheer (Hrsg.). Heft 89, Saarbrücken 1992

[7] Schallar, Ros (2005), „Prozessmanagement in der Bauleitung “, HTW des Saarlandes, Saarbrücken, Diplomarbeit

[8] Schiltz, Yves (2003), „Rechnergestützte Planung und Steuerung von Baustellen“, Universität Luxemburg, Luxemburg, Diplomarbeit

[9] Siegwart, Sven (2004), „Umsetzung von Controllingwerkzeugen in klein- u. mittelständischen Betrieben“, HTW des Saarlandes, Saarbrücken, Diplomarbeit

[10] Thomas, Steffen (2004), „Prozessmanagement im Ingenieurbüro“, HTW des Saarlandes, Saarbrücken, Diplomarbeit

Basic Study of Smart Robotic Construction Lift For Increasing Resource Lifting Efficiency in High-Rise Building Construction

Chang-Yeon, Cho[1]; Soon-Wook, Kwon[2]; Junbok Lee[3], Suk-Jun,You[4] Sang-Yoon-Chin[5] and Yea-Sang, Kim[6]

[1] Dept. of Architecture & Civil Engineering, Sungkyunkwan University,Suwon-si, Korea, P.O. Box 440-746, PH(82)31-299-4772 ; FAX(82)31-290-7570 ; e-mail :yms09d@hanmail.com
[2] Dept. of Architecture & Civil Engineering, Sungkyunkwan University,Suwon-si, Korea, P.O. Box 440-746, PH(82)31-299-4765 ; FAX(82)31-290-7570 ; e-mail :swkwon@skku.edu
[3] College of Architecture & Civil Engineering, Kyung Hee University, Yongin-si, Korea, P.O. Box 449-701, PH(82)31-201-2541 ; FAX(82)31-204-6780 ; e-mail : leejb@khu.ac.kr
[4] Dept. of Architecture & Civil Engineering, Sungkyunkwan University,Suwon-si, Korea, P.O. Box 440-746, PH(82)31-299-4772 ; FAX(82)31-290-7570 ; e-mail : seokjoon.you@gmail.com
5 Dept. of Architecture & Civil Engineering, Sungkyunkwan University,Suwon-si, Korea, P.O. Box 440-746, PH(82)31-290-7568 ; FAX(82)31-290-7570 ; e-mail : schin@skku.edu
[6] Dept. of Architecture & Civil Engineering, Sungkyunkwan University,Suwon-si, Korea, P.O. Box 440-746, PH(82)31-290-7561 ; FAX(82)31-290-7570 ; e-mail :yeakim@skku.edu

Abstract

Most high-rise buildings are erected in downtown areas, where on-site storage space for construction materials is typically insufficient due to limited spaces. Just-In-Time (or JIT) concept has been adopted so that necessary stock materials and storage spaces can be reduced; for high-rise buildings, however, transporting materials vertically using lifts still poses significant efficiency because its efficiency drops exponentially as the building height grows.

Most of current efforts to counter this problem mainly focus on developing smart tower cranes, whereas lift has gathered less attention in this respect.

This research aims to develop a robotic lift capable of autonomous operations at night time, based on the intelligent lift development toolkit which was previously developed by the authors. It is now in a preliminary study phase.

This study proposes a concept model of the robotic lift system and its operation plans, utilizing several technologies such as an optimized material dispatch algorithm and ubiquitous sensor networks (USNs). The proposed system can be relayed to the horizontal transportation robots located in each floor to move the lifted materials to their destinations.

The outcome from this study will contribute to the improvement of the overall efficiency of the high-rise building construction logistics and space constraints of the construction site. It is also expected that the mechanical performance of the existing lift system can be benefited from this research.

When overall system development, including optimized operation planning model and monitoring subsystem, is finished, it will contribute to innovation of the construction technology.

Keywords: Smart Lift, Robotic Construction Lift, High-rise building construction, operation optimization, concept model

1. Introduction and Motivation

According to 'Skyscraper' webzine, around the globe there are so many ongoing highrise building projects, and competition among the new economic giants such as China and Dubai for the tallest building is apparent. The examples include Chicago Spiral building (150 floors, 610m), Ablaza Al Bite building (76 floors, 595m), Russia's Federation Tower (93 floors, 506m), World Finance Center in Shanghai (101 floors, 492m), etc.

According to an earlier research (Lee et al. 2008), height of a building in construction affects efficiency of lifting equipments installed on site, which also affects the overall schedule. Current generation of the lifts fall short of meeting the field needs, because they aren’t fast enough (typically 100-150m/min) to elevate materials to the desired floor on time if the floor is located above 300m; as a result, overall schedule is significantly affected from it.

To counter the efficiency issue, various techniques from working overtime to Six Sigma have been applied, only to reveal other issues such as workplace safety and optimization problems due to spatial and temporal constraints.

This research aims to develop fundamental concepts for next generation lift systems which allow higher efficiency and better management using intelligent control. To achieve this goal, the authors have analyzed existing lift systems with respect to their feasibility for high-rise application.

2. Construction Lift (Literature Reviews)

2.1 Classification of the lift systems

There are several different lift types commonly used in construction sites: first, construction hoists which move materials and men vertically; second, simpler ones which resemble passenger elevators and used for smaller loads; third, hydraulic ones. Construction lifts may be further sub-categorized into a low speed type, a medium speed type, a high speed type; a high capacity type, and special types. These types can be briefly described as follows:

a. Low-speed type: most commonly used in construction sites. Suitable for lower-rise buildings such as community housing and office buildings. Typical speed for this type is 38m/min, and maximum lifting height is typically 150m; also, its nominal load is 1.0 ~ 1.2 tons.

b. Medium-speed type: Suitable for medium-to-high rise projects. Hoists of this type generally have better loading capacity than the low-speed ones. Their maximum installable hight is 150m to 300m, and their nominal load is 1.5 tons to 2.0 tons.

c. High capacity type: designed to increase lifting efficiency in mid-to-high rise projects. Its speed is virtually same as the low speed type, but can be installed to higher places (maximum installation height: 150m to 200m), with increased nominal load (1.2 tons to 2.0 tons).

d. High speed type: designed specifically for high-rise projects. Its speed is 100m/min, maximum installation height is at least 350 meters, and nominal load is 2.5 tons to 3.0 tons.

Table 1. Summarizes the performance of these lift types. Seen in the table, their mechanical performance differences come from the use of inverters, and location of the drive motors.

types	Nominal loads (ton)	Lifting speed (m/min)	dimension of the loading room (m)	Max. size of the door opening(m)	Frequency of power alternation (Hz)	Nominal power consumption (KVA)	Location of the drive motors	Use of inverters
Low speed	1.0-1.2	38	1.27x2.9x2.5	1.27x2.4	50/60	2x27KVA	internal	no
Med speed	1.2-1.5	70	1.5x4.0x2.65	1.5x2.6	50/60	2x47KVA	rooftop	yes
High capacity	1.2-2.0	38	1.5x4.0x2.65	1.5x2.6	50/60	2x75KVA	rooftop	yes
High speed	2.0-3.0	100	1.5x4.5x2.65	1.5x2.6	50/60	2x64KVA	rooftop	yes

2.2 Mechanical performance of construction hoists and survey of existing

Figure 1 illustrates the nomenclature of a typical construction hoist. It is mainly composed of a loading cage (which actually move people and materials vertically), a mast which sustains the cage, a rack-and-pinion gear assembly which converts rotational force to vertical movement (for driving the cage), and a counterweight system. Details of the drive train are illustrated in Figure 2.

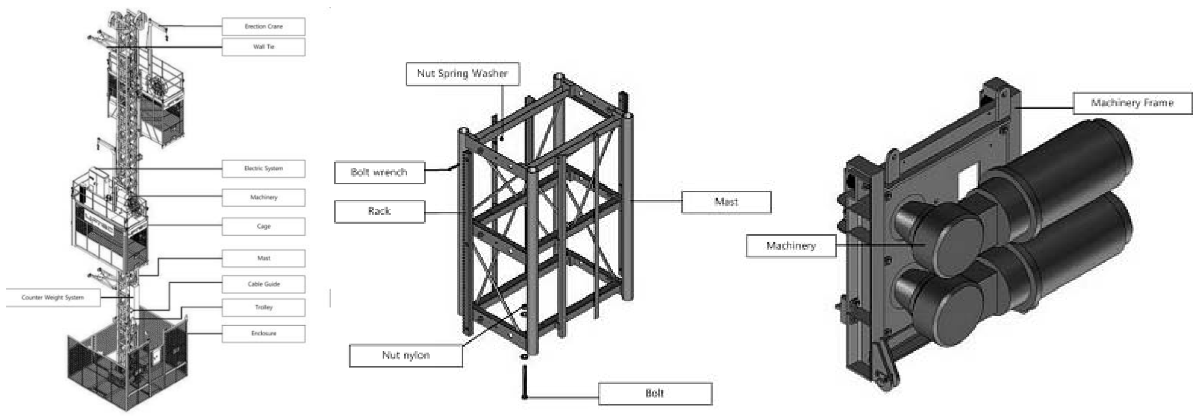


Figure 1. Typical Construction Lifts and Details of the Drive Train

A hoist can determine where it is located by counting teeth of the pinion gear while it is rotating, which can be adjusted with magnetic markers placed in each floor.

2.3 Survey of precedent researches

For the authors, it was unable to find any known researches in (at least in Journals of ASCE and Automation in Construction- the search was by no means exhaustive, as some undisclosed technical reports or regional research articles might have been omitted); on the other hand, most researches for automated material movement focused on tower cranes and construction robots. Majority of domestic researches in South Korea appeared to discuss management techniques for lifting efficiency rather than development of new equipments. They are summarized as follows:

(Lee JB and Han CH, 2008) evaluated economic feasibility of using medium speed hoists in buildings of 30-40 floors. (Chung et. al, 2004) suggested pre-fabrication and carefully planned load-balancing for lessening hoist workloads. In (Lee et. al, 2004), as a part of enterprise resource planning, improved hoist scheduling was proposed to implement just-in-time delivery. (Kim and Han, 2008) proposed a processof determining optimum number of hoists needed in systematic manner. (Park et. al, 2001) discussed optimization process of the hoist schedules for finish works.

3. As-is Problem Statements

3.1 Analysis of of daily hoist travels with respect to lifting height

In our previous research, efficiency of a hoist can be measured by the time needed to lift things to the designated height, whose approximate value can be calculated using the following expression:

Time needed to lift = {average lifting height (m) / lifting speed (m/min) x2 } + lead time (for loading and unloading)

Table 2 shows the number of estimated single lift times and total number of daily travels of different hoist types if they were installed in Burj Al Arab Hotel project, which is 321 meters high. In this table, daiy work time is estimated to 8 hours, and Machines are operated in 70% to their full capability.

Table 2. The number of estimated single lift times and total number of daily travels of different hoist types

Lifting height (m)	Lead time (min)	Low speed type		Medium speed type		High speed type	
		Time of single lift (min)	Number of daily travels (travels/day)	Time of single lift (min)	Number of daily travels (travels/day)	Time of single lift (min)	Number of daily travels (travels/day)
20	10	11.05	31	10.57	32	32.00	33
40	10	12.11	28	11.14	31	31.00	32
60	10	13.16	26	11.71	29	29.00	31
80	10	14.21	24	12.29	28	28.00	30
100	10	15.26	23	12.86	27	27.00	29
120	10	16.32	21	13.43	26	26.00	28
140	10	17.37	20	14.00	24	24.00	28
160	10	18.42	19	14.57	24	24.00	27
180	10	19.47	18	15.14	23	23.00	26
200	10	20.53	17	15.71	22	22.00	26
220	10	21.58	16	16.29	21	21.00	25
240	10	22.63	15	16.86	20	20.00	24
260	10	23.68	15	17.43	20	20.00	24
280	10	24.74	14	18.00	19	19.00	23
300	10	25.79	14	18.57	19	19.00	23
320	10	26.84	13	19.14	18	18.00	22

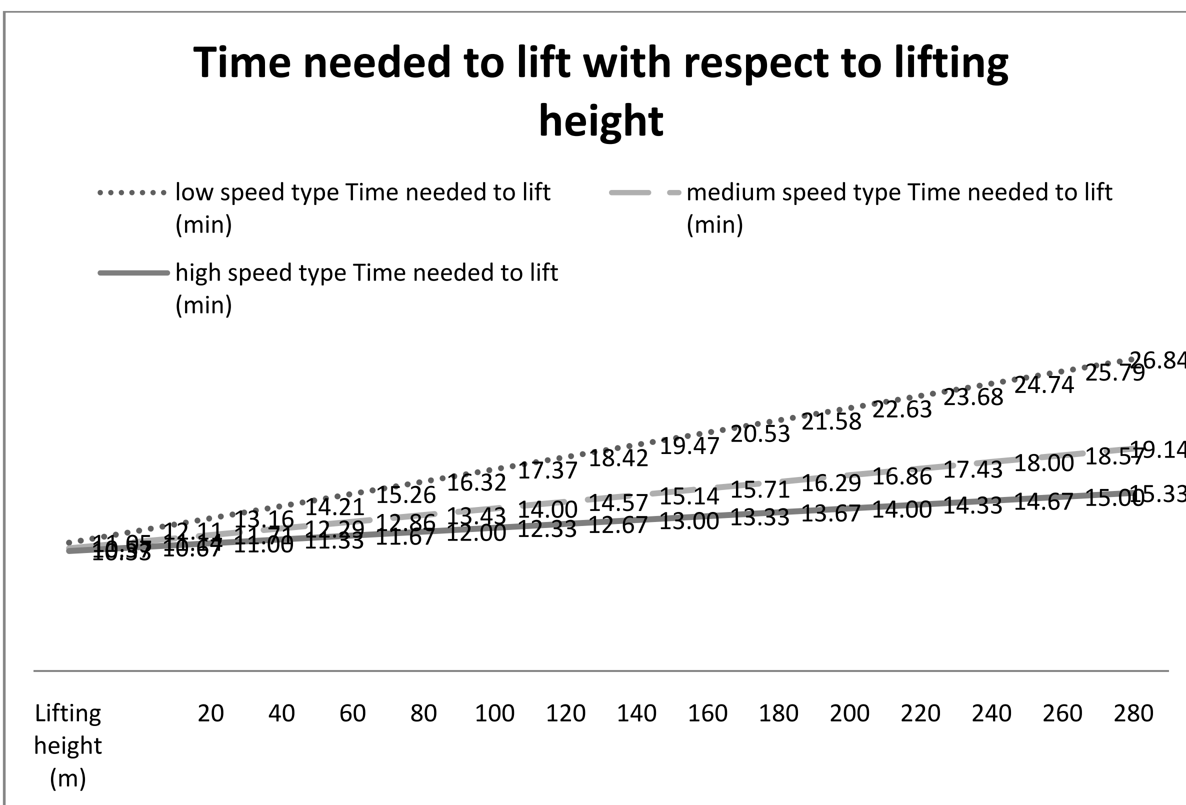


Figure 2. Time needed to lift with respect to lifting height

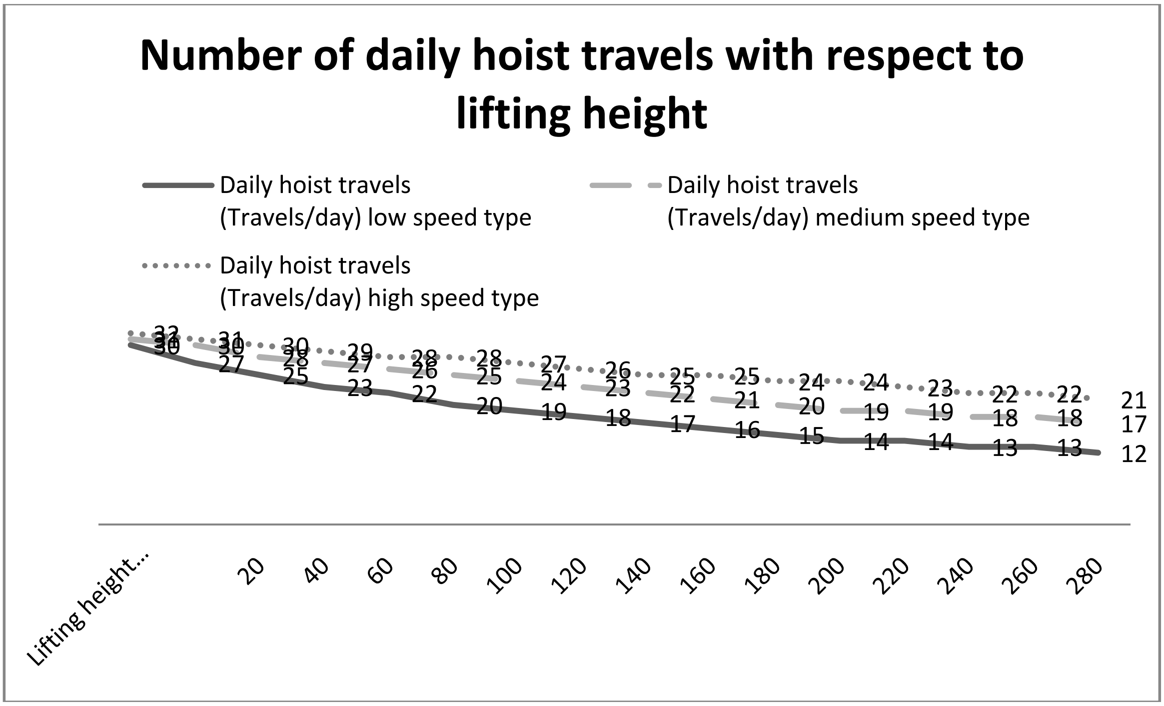


Figure 3. Number of daily hoist travels with respect to lifting height

According to Table 2, Figure 2 and Figure 3, number of hoist travels decreases as building height increases. For Burj Al Arab (321m), its daily travel would be limited to 22. For high-speed lifts

For different lead times ranging from 5 to 10 minutes, Table 3 and Figure 4 shows their effects.

Table 3. Sensitive Analysis For different lead times ranging from 5 to 10 minutes

Lifting height	leadtime 10min	leadtime 9min	Leadtime 8min	leadtime 7min	leadtime 6min	leadtime 5min
20	32.52	36.00	40.32	45.82	53.05	63.00
40	31.50	34.76	38.77	43.83	50.40	59.29
60	30.55	33.60	37.33	42.00	48.00	56.00
80	29.65	32.52	36.00	40.32	45.82	53.05
100	28.80	31.50	34.76	38.77	43.83	50.40
120	28.00	30.55	33.60	37.33	42.00	48.00
140	27.24	29.65	32.52	36.00	40.32	45.82
160	26.53	28.80	31.50	34.76	38.77	43.83
180	25.85	28.00	30.55	33.60	37.33	42.00
200	25.20	27.24	29.65	32.52	36.00	40.32
220	24.59	26.53	28.80	31.50	34.76	38.77
240	24.00	25.85	28.00	30.55	33.60	37.33
260	23.44	25.20	27.24	29.65	32.52	36.00
280	22.91	24.59	26.53	28.80	31.50	34.76
300	22.40	24.00	25.85	28.00	30.55	33.60
320	21.91	23.44	25.20	27.24	29.65	32.52

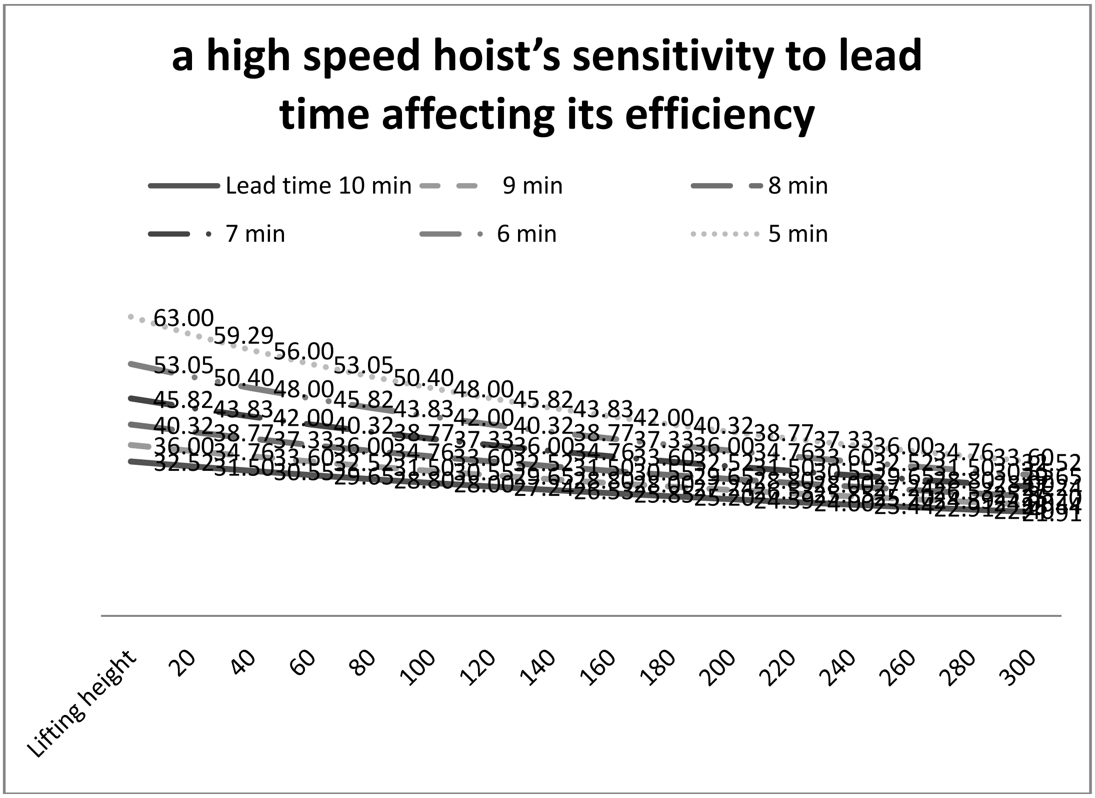


Figure 4. a high speed hoist's sensitivity to lead time affecting its efficiency

To improve the efficiency issues, substantial changes in existing hoist mechanism and their shape are obviously desirable; however, for seeking ready-to-use options, shrinking the lead time to the minimum seems to be a reasonable approach.

3.2 Analysis of existing work process

To understand the lead time in detail, we attempted to analyze the existing work process. We divide the workers related to hoist operation into two groups; the one for ground-level crews loading the hoist with materials, and the other for destination-level crews unloading the delivered materials. Figure 5 is a flow diagram of the work process.

In this figure, workflows of two work crews (ground-level crews and destination level ones) are clearly identified. At the ground-level floor, the crews would use the entrance door where as the destination crews would use the exit door.

The ten-minute lead time assumed in Table 2 corresponds to the 'stand-by' time in Figure 6, which implies that reduced crew-work would positively affect the stand-by time. Further breakdown of the workflow for both crew groups (and their time consumption) can be drawn as Figure 6.

Let Tc1 (stand for Check Time) for the time needed for ground crew to identify total amount of the material to be loaded, Tl for the time needed to load the hoist. The stand-by time is then described as Tc1 + Tl. Likewise, Let Tc2 and Tu for the check time and time unloading time of the destination floor crews respectively. The stand-by time for the destination floor crew is described as Tc2 + Tu also.

Therefore, Total standby time (denoted as Ts) in a single lift cycle can be summarized in the following expression. Our intelligent hoist concept intends to reduce Ts, which will be further described in following chapters.

4. Automation Technology Analysis

4.1 Available technologies

Barcode systems have been used for supply chain management, which is being replaced by newer RFID systems. Bar code readers have shorter recognition ranges, with inability to scan multiple tags at a time, whereas RFIDs have longer ranges and simultaneous scanning capability.

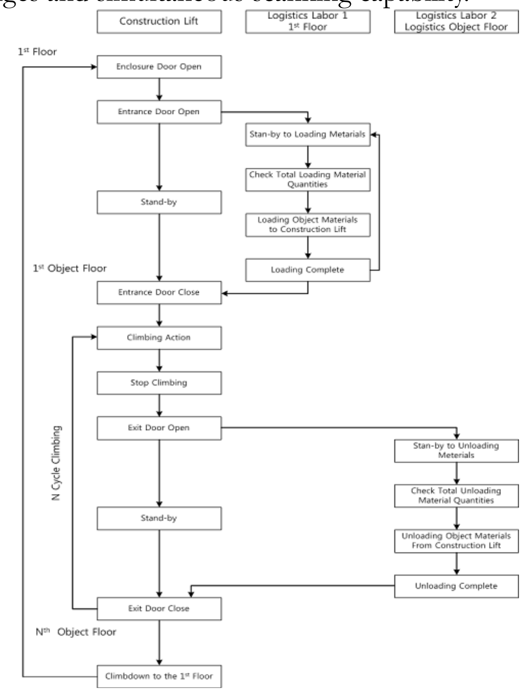


Figure 5. Flow diagram of As-is work process

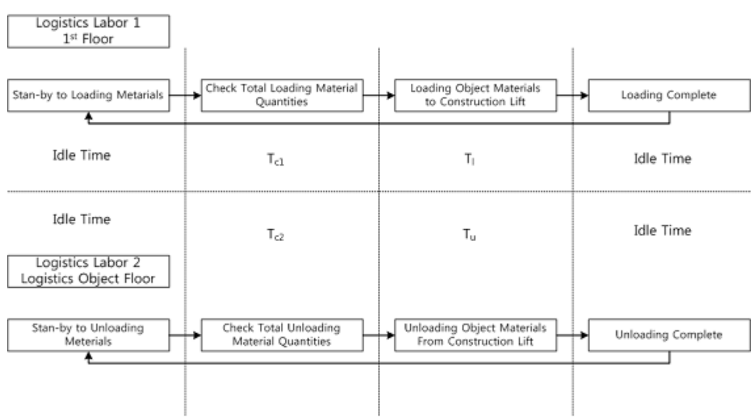


Figure 6. the workflow for both crew groups

In South Korean research society, RFIDs are explored with USN (Ubiquitous Sensor Network) for managing logistics in construction sites.

Among the listed researches, in the 'Next generation intelligent construction logistics automation system', now in its 3rd year, a toolkit for RFID material recognition, which is applicable to hoists as well enabling automatic request of the materials from each floor to the sensor-equipped hoist, is under development. The research also proved that the RFID-based system can interact with real-time construction project management system. The following Figure 7(Research report: An intelligent construction logistics automation system. Ministry of land, transportation, and maritime affairs) is a summary of this research.

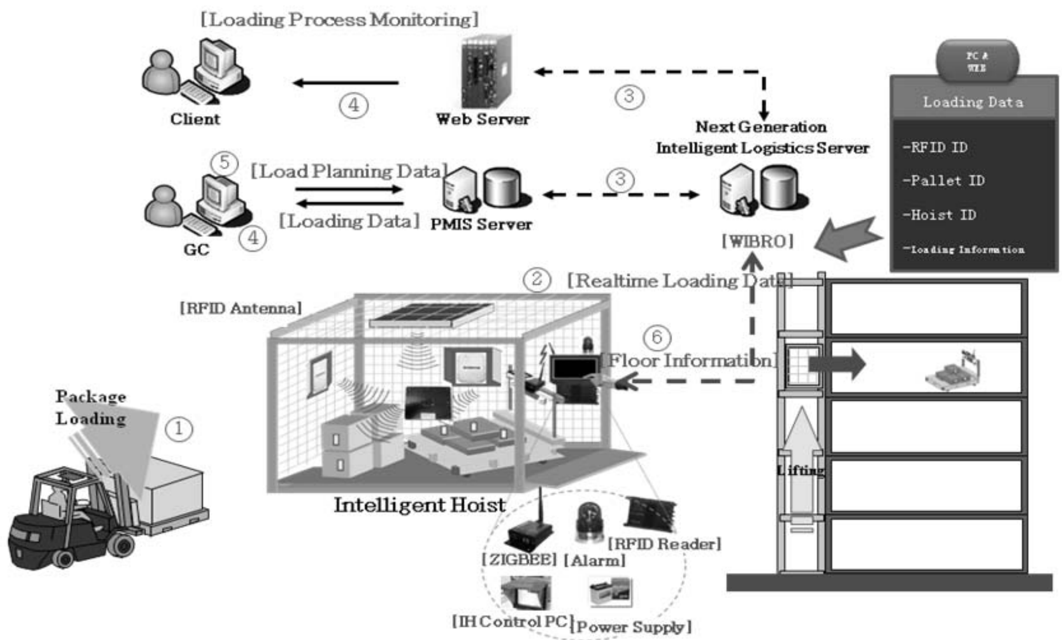


Figure 7. Development of Intelligent Construction Logistics System

Judging from these research cases, RFID is deemed to be feasible for replacing human labor regarding to checking materials prior to loading and unloading (i.e., Tc), thus reducing the check time.

First of all, the process of selecting the optimum tower cranes for super tall building projects such as collection of actual data, application of the data, decision-making support, etc. was delineated. Then, the function and logic for the development of the system for the selection of optimum tower cranes were analyzed, and the method for the application of the results of this study to the development of such a system was investigated.

4.2 Smart Construction Lift Requirement for High Rise Building Construction

Resulting from the research surveys and the work process study, we have defined the required functionalities of the smart construction lift for high-rise buildings, they are:

First, automated recognition of the destination floor for each material loaded to the hoist (with help of PMIS information)

Second, identifying material demands from each unloading floors

Third, performance for reducing loading/unloading time

We also laid out the possible approaches for implementing these requirements: The first requirement can be fulfilled with the development of a PMIS communication module. The second one is realized by RFID systems for sensing what the hoist currently has, besides by the PMIS communication for what needs to be unloaded to each floor. The third (and the last) one is possible by devising a new self-loadable end effector. The work process enhancements benefited from this projected development is illustrated in Figure 8.

5. Conceptual Design

5.1 Conceptual design of Smart Construction Lift

The smart construction lift is composed of the following modules, targeted for shorter stand-bys.

First, a self-loading module

Second, an RFID tag scanning module capable of identifying the loaded materials and their quantities.

Third, a communication module to get destination information (e.g. destination floor of given materials and their quantities) from PMIS.

Fourth, active floor recognition module for ever-varying floor information.

Fifth, a calculation module for travel optimization to a target floor

These modules and their functions are listed in the following figure.

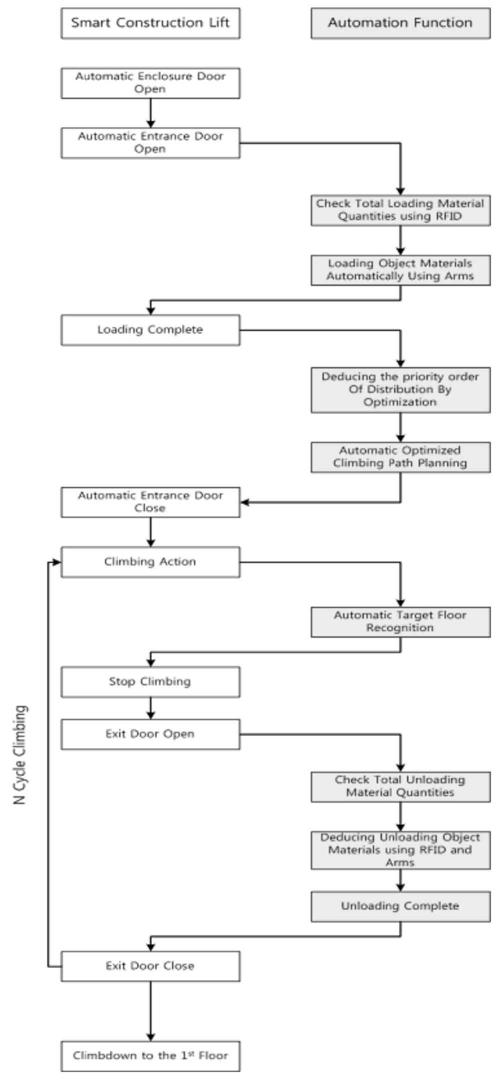


Figure 8. Work process enhancements benefited from this projected development

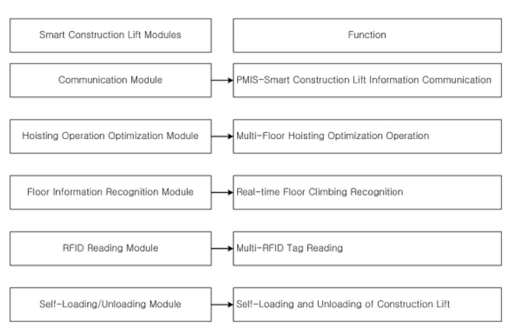


Figure 9. Smart Construction Lift Modules

The data flow between the modules and the tasks performed by each module is illustrated in Figure 11.

5.2 Proposed conceptual design for Smart Construction Lift

For each module in Smart Construction Lift, they can be realized using the following technologies.

1. Communication modules can be realized using wireless communication technologies such as Mobile WiMax, WiFi, or Cellular network.
2. RFID modules can be implemented using the RFID readers and antennae
3. Floor recognition module can be implemented with an encoder, a pinion gear counter, and

magnetic sensors.

4. Hoisting operation optimization module requires a control computer to calculate the optimized travel plan, which can be served with a personal computer

5. Self loading/unloading module can be implemented with an active hydraulic arm or a fork

The configuration of the proposed modules is illustrated in Figure 12, forming a conceptual design of the Smart Construction Lift.

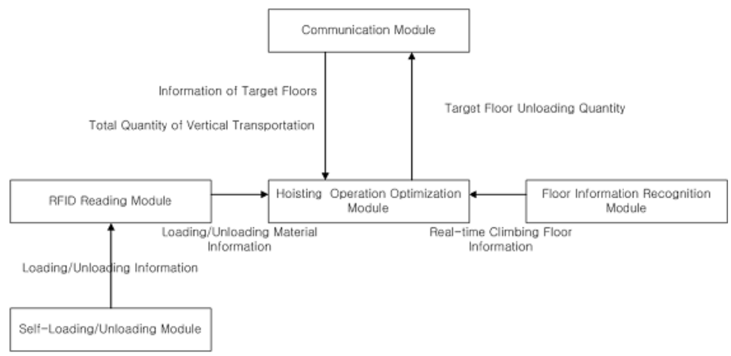


Figure 11. data flow between the modules and the tasks performed by each module

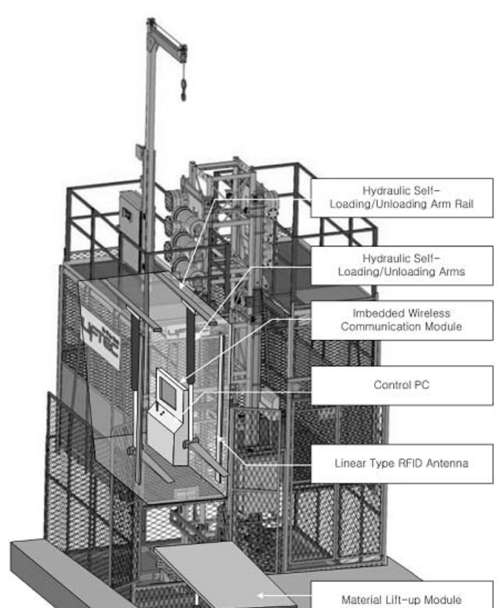


Figure 12. Conceptual Design of Smart Robotic Construction Lift

For implementing the hydraulic loading/unloading module, an fork type would be better if materials are conveyed on pallets and/or their quantities are big; for individual materials without pallets, arm types would be better.

For wireless communication modules, an embedded module (installed inside the control PC) is expected to fare well in terms of noise resistance rather than external one (hooked up via cable).

For RFID antennae, circular types seem better than linear types because construction lifts reflect radio waves causing interference

5.3 Expected Benefits

As mentioned in previous sections, this research aims to propose an alternative to conventional lifts for increased productivity without the need for completely redesigning them. For building projects where vertical transport accounts for significant amount of construction schedule, such benefit would be maximized.

The expected benefits of the proposed smart lift system can be summarized as Table 5, which is shown below.

Table 4. as-is logistic management matrix

	Labor 1	Labor 2	Lift Operator	Lift
loading packages at Storage Area				
moving packages to lift entrance				
Waiting for loading				
loading to lift				
Input target floor				
Lifting packages				
Fit lift-exit hight to floor hight				
Open lift-exit				
unloading				
close lift-exit				
lifting next target floor				
Fit lift-exit hight to floor hight				
Open lift-exit				
unloading				
close lift-exit				

Table 5. To-be logistic management matrix

	Labor 1	Labor 2	Lift Operator	Lift
loading packages at Storage Area				
moving packages to lift entrance				
Waiting for loading				
loading to lift	1			
Input target floor	2			
Lifting packages				
Fit lift-exit hight to floor hight				3
Open lift-exit			Not Needed	
unloading				
close lift-exit				
lifting next target floor				
Fit lift-exit hight to floor hight				3
Open lift-exit				
unloading				
close lift-exit				

- 1 : automated loading system using hydraulic arms or folks
- 2 : automated system using RFID
- 3 :automated floor sensing using light sensor and encoder

A to-be-developed hydraulic arms and forks will enable automatic loading and unloading (thus reducing labors needed for those operations). Also, overall lead time required prior to actual lifting can be reduced by various measures such as: utilizing RFID-based automatic logging, automatic identification of the destination floor of a given package with help of onboard control computer, identification of the lift's current location which will be implemented using light sensors and encoders. Such improvements will allow unmanned operation that won't require human operators, and they will enhance productivity of the storage-area operations thanks to shortened lead time.

6. Conclusions and Further Studies

The authors have proposed a concept of a smart construction lift, whose functionality can be summarized as follows:

First, automatic loading/unloading capability which enables unmanned operation.

Second, wireless communication with PMIS

Third, optimization capability in travel planning with help of PMIS information and automatic operation of the lift using the optimized plan

Fourth, automatic operation to the target floor.

Fifth, automatic unloading upon arrival at the target floor

Sixth, sending information about unloaded materials to PMIS in real time.

Once realized, the smart construction lift will help logistics in high-rise construction sites, where sending materials to higher locations often cause overall construction schedule delayed due to mechanical limits of the lifts. Further research plans include the prototype development and evaluation of it. Also, this research will contribute to the construction industry by demonstrating a cost-saving alternative to the revolutionary approaches to completely replace the existing rack-and-pinion mechanisms, which require substantially more money and time, besides the cost of 'trial-and-errors'.

References

[1] Research report: An intelligent construction logistics automation system. Ministry of land, transportation, and maritime affairs (Republic of Korea).

[2] Lee, J.B. et. al.(2008) A Feasibility Study on Optimal Lifting Planning in the High-rise Apartment Building Construction, Journal of Korea Institute of Construction Engineering and Management Institute 9(3), pp. 185-193 (Korean)

[3] Ho, J.K. (2007). A System for the Selection fo the Optimum Tower Cranes (Opt-TC). Journals of KICEM. 8(6). (Korean)

[4] Kim, A.Y. (2007). A Basic Study on Selection of Climbing-type Crane and for Structuring the Stability Analysis Model. Proceeding of 13th Korean Institute of Ecological Architecture annual conference. (Korean)
[5] Kim, J.J. (2005). A Study on the Hoisting Planning System in Highrise Building Construction. Journal of Korea Institute of Construction 5(4). (Korean)
[6] Chung, Y.W. (2004). A study of the Highrise Building's Lift-up Management: Through Division of Loading Factors of Equipment Materials. Proceedings of 5th Annual Conference of KICEM. (Korean)
[7] Lee, H.S. (2004). A Study on the Method of New Activity Based Cost Management Coping With Changes in the Cost Structure of Real Estate Construction Industry. Journal of KICEM 5(1). (Korean)
[8] Kim, S.K. (2008). A Study on the Estimation of Proper Numbers of Construction Lifts. Journal of KIC 8(3). (Korean)
[9] Liftec Construction Lift Installation Manual. Pp.25. Liftec Inc, 2008.8

Work State Identification using Primitive Static States –Implementation to Demolition Work in Double-Front Work Machines-

Mitsuhiro Kamezaki [1], Hiroyasu Iwata [2] and Shigeki Sugano [3]

[1]Department of Modern Mechanical Engineering, Waseda University, Shinjuku-ku, Tokyo, Japan, kame-mitsu@sugano.mech.waseda.ac.jp
[2] Waseda Institute of Advanced Study, Waseda University, Shinjuku-ku, Tokyo, Japan, jubi@waseda.jp
[3] Department of Modern Mechanical Engineering, Waseda University, Shinjuku-ku, Tokyo, Japan, sugano@waseda.jp

Abstract

Double-front construction machinery, which has been designed for adaptation to complicated work, demands higher operational skills to control two manipulators with more multiple joints. To handle more complicated machinery skillfully, intelligent systems that can autonomously identify the current work states and also provide cognitive and operational supports to their operators are inevitably required. Particularly, work state identification methods strongly require high reliability and robustness due to the variety of the construction work environment and operator's skill level. However, most current construction machinery has unique functions that only reproduce the movements originating from the operator. We therefore addressed the need for a new conceptual design of operator support system and evaluated it using our newly developed simulator. Our experimental results showed that the support system improves the work performance, including decreasing the operational time for completing a task, the number of error operations, and the mental workload on the operators.

Keywords: Construction machinery, Intelligent system, State identification, Operator support.

Introduction

The adaptation of construction machinery to highly skilled, complicated work has been expected. Such tasks, which include sorted dismantling for recycling and reusing resources, rescue and recovery work at a disaster site and building construction, are different from the conventional simple earthwork tasks such as ground levelling, transportation, excavation, and loading. In response to such changes in recent social needs, double-front construction machinery (DFCM), which has two manipulators, was developed, as shown in the right side of Figure 1 (detailed specifications are given in (Ishii, A. 2006)).

In conventional single-front construction machinery (SFCM) such as excavators, breakers, or cranes, one manipulator is operated using two control levers. On the other hand, DFCM has two manipulators operated using two control levers (Figure 1). When comparing them with SFCM, although the adaptability to a wider range of construction works is surely improved, the manipulators have more than twice the number of degrees of freedom, and therefore, operators need extremely high level of operating skills. This is a major drawback that can lead to lower efficiency and work quality by making machine operations more confusing. Additionally, it could lead to dangers being overlooked, such as operators not noticing the existence of outside workers or warnings from co-workers, because operators are concentrating more on the difficult machine operations.

As one means of addressing of these skill and safety problems, we suggest an intelligent system that makes it easier for an operator to operate a machine and to accomplish complicated tasks. This type of intelligent system framework is composed of the functions to autonomously distinguish between several states such as a working or a dangerous state, and to provide information and operational support corresponding to the identification results. For example, in minute work in which an operator endures a lot of stress, such as moving breakable things using double-front, the machine provides an automatic switching support of the operation gain, or in a dangerous state, such as those overlooked because of an operator's devotion to a machine operation, the machine informs the operator of the danger.

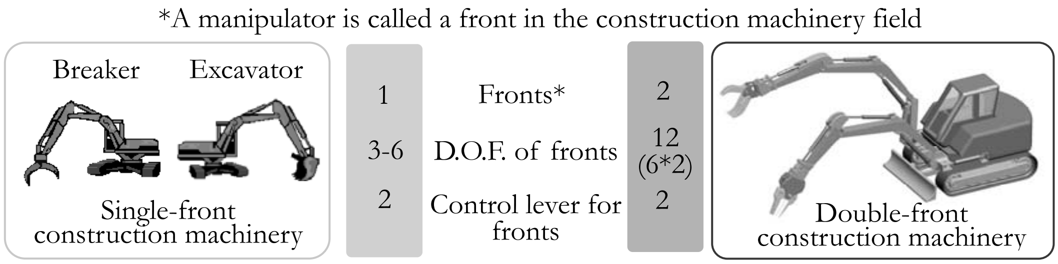


Figure 1. Greater difficulty in controlling advanced construction machineries.

However, most of the current construction machinery has unique functions that only reproduce the movements originating from the operator. Even an angle sensor to detect the posture of a manipulator is rarely used. Technology for the intelligent control of construction machinery has conventionally been developed in an application-specific way, and research efforts have been devoted to areas such as oscillation-stopping control for cranes (Takemoto, A. 2004), (Yoneda, M. 1997), remote operation of excavators (Sakai, R. 2004), intelligent oil-hydraulic control (Sakurai, Y. 2004), and analysis of power shovel operation (Rodriguez, J. 2004). However, comparing the fields of advanced automobile (Inoue, G. 2008) and surgical robot (Sun, L. W. 2007), which have human-operated machine systems like DFCM, there has been less movement toward incorporating an intelligent system.

We thus aimed at constructing a new framework for an intelligent system for construction machinery. To make it easier to adapt this system to fit all construction machinery, we targeted DFCM, which uses the most complicated hardware constitution at present. On the basis of above, we have designed an operator support system for DFCM.

Analysis of Intelligent Support Scene

It is important to adequately consider the characteristics and problems associated with construction machinery and construction sites. In this chapter, we derive the overt or covert needs, and embody a highly effective intelligent support scene.

Fundamental Problems

We analyzed the actual condition and problems with construction machinery and construction sites. We wrote them up in the following elements after reviewing construction documentations and talking to construction company workers.

1) Construction machinery: a) The size and weight are big, so the risk of a collision with a worker is extremely big, b) there is a delay in the oil pressure system and the inertia force is big, so instantaneous or quick movements are difficult and dangerous, c) there are many blind spots for an operator, so it is difficult for them to check the danger spots, and d) the vibration and noise created by the machine's movement decreases the cognitive ability of an operator and promotes fatigue.

2) Construction sites: a) The construction machinery and outside workers must cooperate, so there is always danger of contact with workers, b) an operator must take excessive care to recognize other equipment (e.g., wheel loaders, dump trucks, or outside workers), so the operator's awareness of the whole situation may weaken, making it hard to concentrate on work.

From these analyses, we are better able to understand that an operator must have the operation skills and cognitive ability to ensure a safe work environment by always grasping the situation while also paying attention to the machine characteristics.

Advanced problems

A standard intelligent support system consists of three modules, an inside or outside information detection module, a state discrimination module for recognizing the working or danger states based on the information detection module, and an operator support module based on the state identification module. Based on the analyses mentioned in the previous section, we collected information on the problems associated with applying this type of intelligent system framework.

1) Information detection: a) Usage environment is inferior, so an indestructible structure and high noise immunity are required for sensors. b) Sensors are not easily mounted on current construction machinery, so a mechanical design taking into account their installation is necessary.

2) State identification: a) The work environment is complicated, and this is especially true for demolition work where the use of DFCM is necessary. b) The shape and position of the objects manipulated in construction work continually change. c) Skill levels and operational methods differ from one operator to another. Therefore, from these three characteristics we have determined that it is much more difficult to adequately identify working states.

3) Operator support: a) The level of demanded support differs depending on the situation and operator. b) Machine characteristic (e.g., inertia or hydraulic system) support is necessary.

We found from these analyses that there are many unknown parameters involved in using construction machinery, so state identification is very difficult. Therefore, we think that it is indispensable to create state identification technology taking into account each characteristics of construction machinery.

Embodiment of Intelligent Support Scene

Based on the analysis results given in previous sections, we extracted the assumed needs, and considered demanded technical standards. We created a specific situation where the support could be useful to the operators and workers, and construction machinery. We used the following procedure to gather feedback for designing the system. Firstly, we asked operators and workers what types of situations they thought some support would be useful, secondly the states that are necessary for concerned support should be provided, thirdly, the required information for state identification and the method involved, and finally the sensors necessary for detection of concerned information. We took into account the input information, working state, support, including the situation, the sensor and its installation position, and then arranged more than 60 situations in which support is necessary. Table 1 shows some examples in which it is thought that the effectiveness is high.

System Design

We now understand that a state identification is a key module of operator support system, which influence the usefulness of the system. We therefore designed an operator support system, particularly focusing a state identification method.

Semi-autonomous Support Method

In a manually human-operated work machine, a fully autonomous system would be the ultimate support method because of having merits, such as not only freeing a machine operator from dangerous work, but also reducing personnel expenses. However, autonomous systems are adapted for use in a limited number of simple tasks where many of the conditions are approximately determined, such as ground levelling or digging a broad mine without on-site workers. When we take into account examples of practical situation use such as the advanced surgical operation support, we can better understand how important it is that the

Table 1. Intelligent support scenes (extract).

SCENE	INPUT	STATE	OPERATOR SUPPORT
Sorted dismantling work	Angle of operation lever	Minute work	Depending on lever operation, machine automatically changes operation gain. In minute work, machine lowers gain. **-Parameter Optimization Support**
Cooperative transporting work with double-front	End-effector position / load	Cooperative transporting work with double-front	Depending on gap of position of end-effectors and load, machine revise relative position of end-effectors **-Operation Support**
Demolition work / Rescue work	End-effector position / object position	High place work /Work at blind side	To feed back image from virtual camera installed in end-effector, machine offers to effectively expand field of vision. **-Information Support**

operator makes the final decisions. Even if a machine autonomously determines that it can directly support the operator in situation recognition or operation, it is important that system only visually and acoustically assist them, and which is especially true for addressing the problems mentioned in Section 2. Therefore, we decided to create a semi-autonomous support method.

State Identification

1) Concept of new method: Many researchers have already reported on different types of state recognition techniques, and the hidden Markov model (HMM (Rabiner, L. R. 1989)), dynamic Bayesian network (DBN (Murphy, K. 2002), (Laet, T. D. 2008)), and support vector machine (SVM (Cristianini, N. 2000)) have been proposed as promising techniques. These methods have the advantages can handle identification systematically by optimization, but for getting a desired output result they must still require an enormous amount of input data for learning, a suitable pre-processing, parameter adjustments, and so on. A systematized theory for a method of adjusting these parameters has yet to be designed, and at present we have only a trial-and-error method.

A state identification method needs to take into account the characteristics of construction machinery and the construction work environment as mentioned in Chapter 2. From the problems mentioned in Section 2.1 1), for points c) and d), an interface provides cognitive support such as reminding the workers what the current situation is that they are working in would be effective. On the other hand, for points a) and b), if construction machinery moves in a direction that an operator did not intend due to a false state identification, the construction machinery can become extremely dangerous to neighbouring workers and environment. On this account, an important point for developing a state identification method is how to avoid misrecognition of work states. Therefore, a work state identification method strongly requires high reliability and robustness that mean not misidentification in any kind of situation. From this standpoint, we understand that it is difficult to use the above mentioned methods (e.g., HMM) that cannot sufficiently and stably respond to the variety of the applied field.

We therefore define a basic work state unit that is completely independent of the various environmental conditions and operator skill levels for certain and robust identification, and that are applicable to all types of construction machinery, including DFCM.

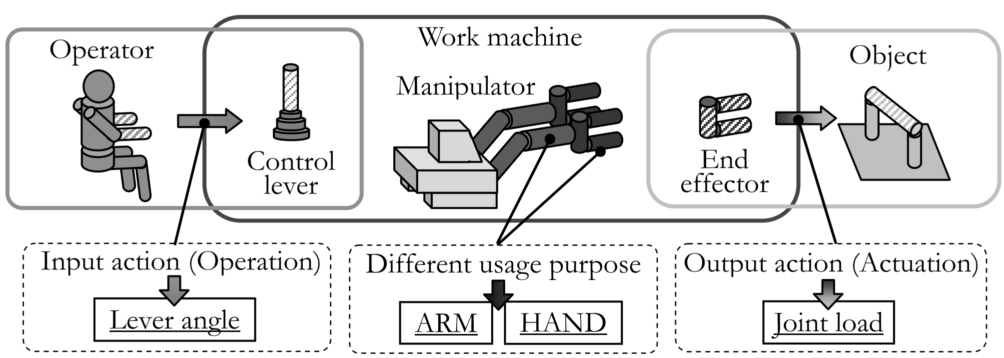


Figure 2. Relations among operator, work machine, and object.

2) Design of new method: As is well known, construction machinery has a human-operated system and applies force to an environment, and therefore, we focused on three corresponding elements: an operator, a work machine and an environment (Figure 2). An operator and a work machine interact through control levers and their interaction represents the existence of the machine operation due to the operator's operation. A work machine and an environment interact through end-effectors and their interaction represents the existence of applied force due to physical contact by machine's actuation. These interactions are defined without using vectorial or time-series information concerning the work object (e.g., position or weight) or the manipulator (e.g., velocity or trajectory), which greatly depends on the work environment condition or operator's skill level, or the machine's specifications, and therefore, the combination of these interaction information can describe the basic work states. From these analyses, we decided to use on-off information for the lever operations and joint load. In addition, focused on the mechanical structure of construction machinery, the attachment part directly interacts with the environment (e.g., grasping or cutting), but the front part is used for positioning the attachment by not interacting with the environment (e.g., reaching). From these different usages, we decided to extract the above information from two parts, the attachment (hereinafter called the HAND), and the front (hereinafter called the ARM).

3) Primitive Static States (PSS): Based on previous sections, we designed a base work state unit using on-off information for the lever operations and joint load, which represents the condition of relations among the operator, work machine and environment, for the HAND and ARM, which represent differences in interaction with the environment either directly or indirectly. These states express the most basic states using static information, so we call these state units primitive static states (PSS). When we focus on a single arm, there are 16 separate states (2^4), and when we focus on for double arm, there are 256 (16^2). In addition, we can apply a specific working state to each of the 16 states (A-P) as shown in Table 2. For example, when HAND load = 0, Hand operation = 0, ARM load = 0, and ARM operation = 1, the PSS(B) is reaching work, and when HAND load = 0, Hand operation = 0, ARM load = 1, and ARM operation = 1, the PSS(D) is compressing work.

Operator Support System

We developed a prototype of an operator support system (OSS) based on the primitive static states. As mentioned in Section 2, the OSS consists of three modules. We concretely explain the contents of these modules. A developed system flow is shown in Figure 3.

1) Information detection: The angle of the control lever and the joint load for HAND and ARM are used as the input information. The former data is obtained from a potentiometer mounted the control lever. As mentioned later, we performed experiments using VR simulator. Therefore, the latter can be easily obtained using the function of the dynamics engine.

2) State identification: From detected information, the PSS identifies the reaching states (B, E and F), grasping state (M and P), and transporting state (L) of each manipulator. We redefined the reaching and grasping states which are demanded to recognize the depth feeling as the long-range-work state and also redefined the grasping states which are demanded precision operations as the minute work states. Additionally, when both manipulators are transporting states, the system identifies this situation as a cooperative transporting state.

Table 2. Primitive static states (PSS).

PSS no.	Input data* HHAA LOLO	Working State (example) (HAND: Grapple)	PSS no.	Input data* HHAA LOLO	Working State (example) (HAND: Grapple)
A-00	**0 0 0 0**	Non-operation and load	I-08	**1 0 0 0**	Holding of object on ground
B-01	**0 0 0 1**	Reaching	J-09	**1 0 0 1**	Abnormal state
C-02	**0 0 1 0**	Holding/ Grasping	K-10	**1 0 1 0**	Holding of aerial object
D-03	**0 0 1 1**	Compressing	L-11	**1 0 1 1**	Transporting/ Bending/ Detaching
E-04	**0 1 0 0**	Hand operation	M-12	**1 1 0 0**	Cutting/ Setting
F-05	**0 1 0 1**	Hand operation in reaching	N-13	**1 1 0 1**	Abnormal state
G-06	**0 1 1 0**	External force during grasping	O-14	**1 1 1 0**	Cutting/ Setting
H-07	**0 1 1 1**	Hand operation in compressing	P-15	**1 1 1 1**	Throwing out

3) Support: In both the minute work and double-front transporting states, the system reduced the operational gain to half in order to make precise work easier. In the long-range-work state, the system presents an enlarged image of the end-effector from a different viewpoint in order to provide positioning in formation assistance as shown in Figure 4 (a small window circled by a broken line at the upper left side of the cab view display).

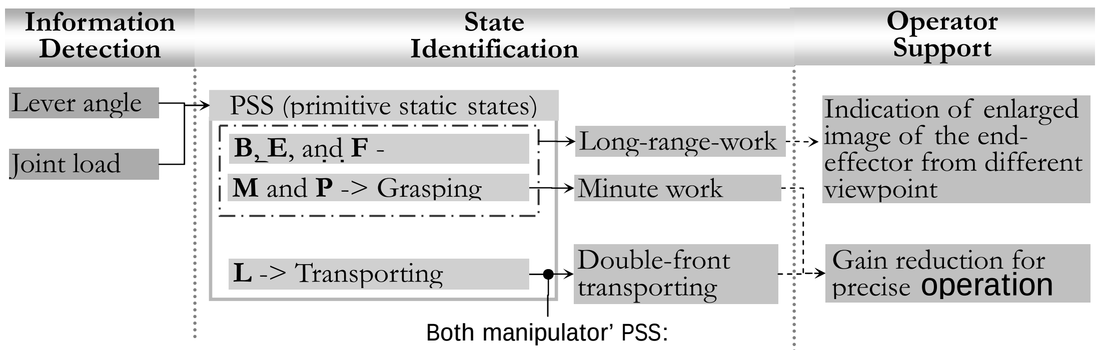


Figure 3. Developed operator support system using primitive static states.

Experiment

We newly developed a double-front construction machinery simulator and performed experiments to inspect the utility of our operator support system using the simulator.

DFCM-VR simulator

To lower the estrangement in the real world, and to simplify the model as much as possible, we implemented only the element that was the most dominant for the construction machinery simulator. The operation lever arranged two joysticks the same as in actual machine. In addition, we reproduced operational gain, sounds, and oil delay, and moreover physical behavior, such as contact judgment functions, inertia, adhesive strength, and frictional force in the environment, which is indispensable in above mentioned work assuming interaction with the environment. We implemented these functions with a physics engine (Open Dynamics Engine (Russell, M)), also with OpenGL and Microsoft MFC. The screen of the graphical user interface and an image from the cab, and the hardware system are shown in Figure 4 (detailed specification given in (Kamezaki, M. 2008)).

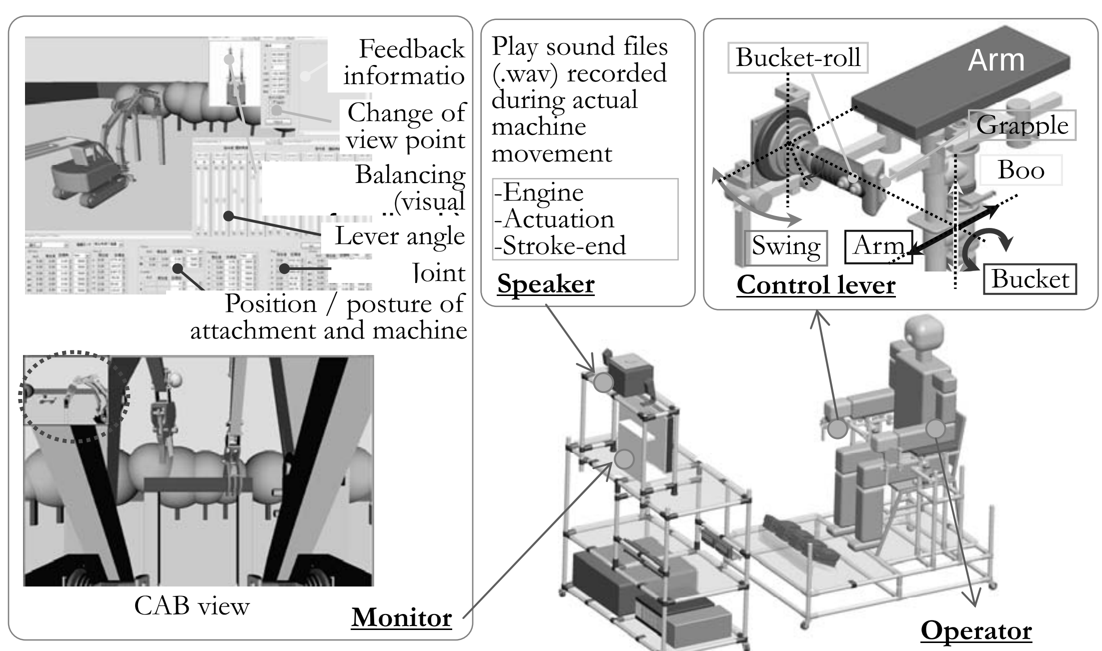


Figure 4. Graphical display and experimental setup of the developed simulator.

Experimental Task

We performed an experiment on minute work and cooperative transporting work states, as shown in Table 1 and Figure 3. We set the removal of a wooden beam as a task. The beam was put on top of two columns in a position to be able to be grasped without the construction machinery being moved. The initial posture of the manipulators was one where both manipulators could be lifted to a high position. An operator reaches just enough to not let the end-effector collide with the beam and grasps it with the two end-effectors. Then the operator places the beam just in front of the two columns. In other words, this task can be divided three states: reaching, grasping, and transportation state. In every state, operation supports are provided as shown in Figure 3. Figure 5 shows the operation supports and work state identification demanded in a work sequence.

Eight healthy adult males (20-25 years old) without any experience at any kind of construction machinery operations were used as subjects. We let them train on the operation for about 20 minutes so that they could get used to the simulator and then perform experiments on three types of patterns in the same task, which were: off-support, support with a gain switch, and support using a view from a virtual camera. All the subjects were randomly tested on all the patterns. We measured the working time (time before grasping from the start and time before placing after grasping). We judged that there was an overload when the relative distance between each end-effector was more than 200 mm. In addition, we quantitatively measure the operator's mental workload by using the NASA-TLX (Hart, S. G. 1988).

Result

The operational time to complete a task is shown in Figure 6, the number of overload in Figure 7, and the mental workload in Figure 8.

We found that the time taken to complete a task decreased with the on-support and also that the virtual camera view support is the most effective. The time taken to complete a task decreased by an average of about 30% when using the virtual camera view, and the maximum-recorded decrease was 70%. We performed two-tailed t-testing about off-support and virtual camera view support, and identified a significant difference $t = 1.85$ ($p < 0.1$). The number of overloads decreased when providing the gain switch and virtual camera view supports. On average there was a 72% decrease (from 6 to 0 times in the best-case scenario). The operator's mental workload decreased when providing supports. We performed two-tailed t-testing about off-support and virtual camera view support, and identified a significant difference $t = 2.79$ ($p < 0.05$).

We found that work performance in more than half of the subjects was improved by our operator support system, and we also found that virtual camera support was the most effective means of support in all evaluations.

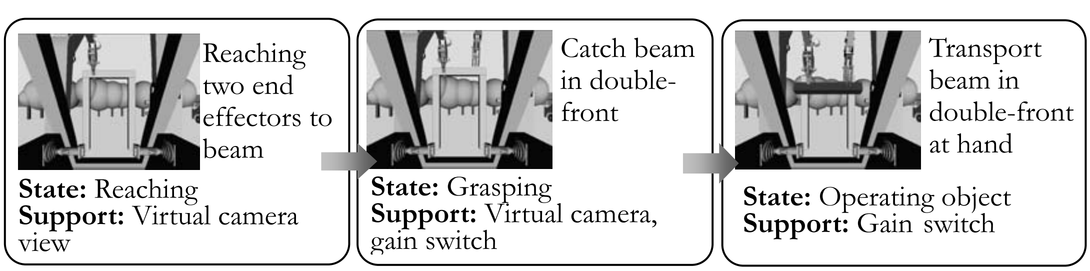


Figure 5. Operation supports demanded in a work sequence.

Discussion

For the inspection of the support effect in the each subject, we calculated the ratio that divided off-support by on-support about above evaluation index about every subject, and Figure 9 shows the results. This figure shows that if the positive value is bigger, the effect of the support was stronger and if the negative value is bigger, the support had an adverse effect.

1) Virtual camera view support: The mental workload when using virtual camera view support is less than two-thirds that of off-support for over half of subjects. We also confirmed that virtual camera view support is effective at both shortening the accomplish time and decreasing the mental workload despite a simply

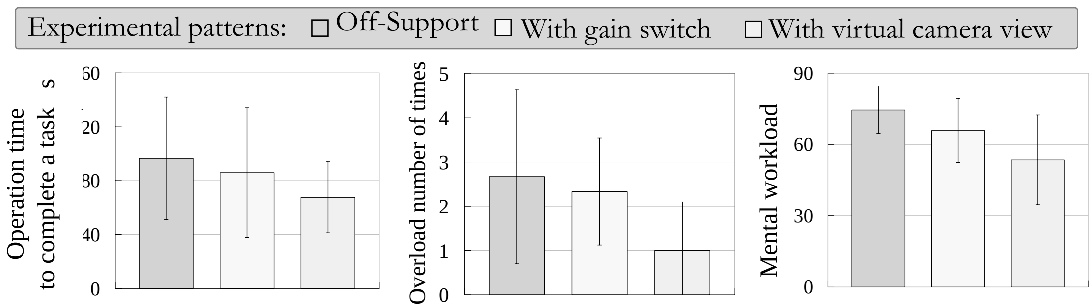


Figure 6. Operation time to complete a task.

Figure 7. Overload number of times.

Figure 8. Mental workload (NASA-TLX).

information support. We think that assistance with depth perception, which is sometimes hard for operators, was effective. Although a positive effect was not found for subjects A and B, it is thought that this support would be more useful if we could control the camera direction and zooming function based on more detailed state identification.

2) Gain Switching Support: The gain switching support provided a decrease in number of overloads for an

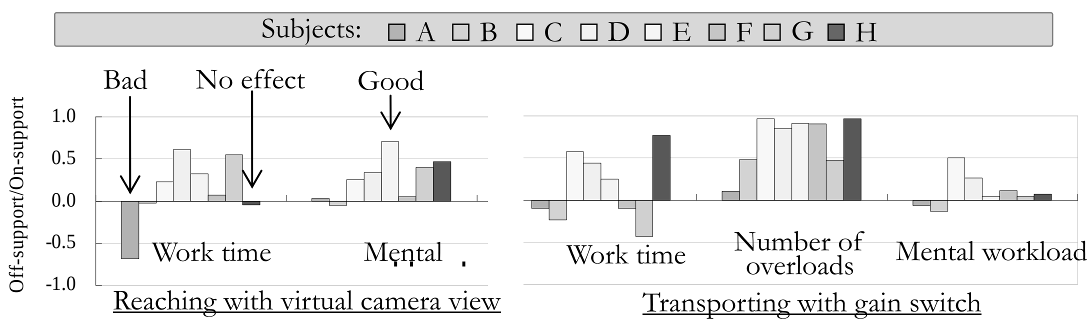


Figure 9. Difference of support effect by subjects.

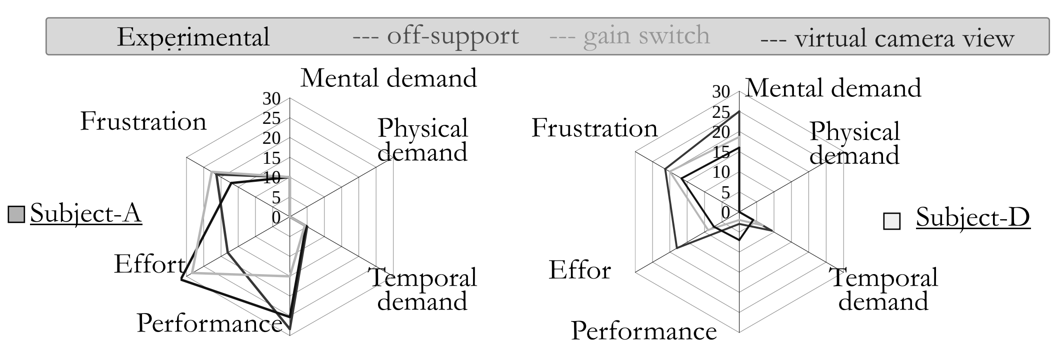


Figure 10. Mental workload - difference of support effect by examinees -.

object. It is essential not to give overload an object during material transportation. We also found that the effectiveness is greatly different for each subject about the accomplish time. Questionnaire results showed that some subjects felt that gain switch support was troublesome when changing the movement properties. For example, mental workload for subjects C, D, and F decrease when using gain switch support (Figure 10), and the EFFORT (an evaluation item of NASA-TLX) particularly shows a remarkable decrease. On the other hand, in subjects A and B, the mental workload increases when using gain switch support. Correspondingly, EFFORT increases by more than 1.5 times. Subject A answered that the movement of a manipulator slowed, so he felt operational effort is increased. From these results, we understood that it was necessary for the gain adjustment method to adequately set the parameter for both an operator's skill level and a task condition.

To provide more useful operator supports, we think that it is important to mutually improve not only the support methodology but also the state identification method. In other words, the state identification level (e.g., number of identifiable states, probability, or static or dynamic) and the operator support type (e.g., cognitive, operational, or parameter optimization) have a close relationship, and there is a suitable identification level for each support. We found that PSS is useful for cognitive support (virtual camera view). For a parameter optimization support (gain switching), we think that the state identification technique using a fixed parameter for support of all situations or variable parameters for9 individual support and combing them is useful.

Conclusion and Future Works

Taking into account peculiar problems for construction work, we proposed the primitive static states which is independent the various environmental conditions and operator skill levels for certain and robust identification. Based on the primitive static states identification, we developed the operator support system, which provides a virtual camera view and gain switch supports for the minute and cooperative transportation works. As a result of having experimented by evaluating novices, we confirmed that the use of on-support decreases the mental workload of operators, the operational time to complete a task, and the number of false operations. Thus, we were able to confirm the effectiveness of the developed operator support system.

For future work, we will present a probabilistic state identification technique, and examine a fusion method with the primitive static states identification. In addition, we will perform an experiment with an actual machine for confirming the practicality of the proposed method.

Acknowledgement

This research was supported in part by Hitachi Construction Machinery Co., Ltd, and in part by Global COE Program in Waseda University “Global Robot Academia”.

References

[1] Cristianini, N., and Shawe-Taylor, J. (2000) “An Introduction to Support Vector Machines. Cambridge.” UK: Cambridge University Press

[2] Hart, S. G., and Staveland, L. E. (1988) “Development of NASA-TLX(task load index) : results of empirical and theoretical research.” in P. A. Hancock & N. Meshkati (eds.), Human Mental Workload, North-Holland, 139-183

[3] Hitachi Construction Machinery, http://www.hitachi-c- m.com (Dec. 2008)

[4] Inoue, G., and Yasunobu, S. (2008) “An intelligent MIMO control and its application to two-wheeled vehicle for driving support.” Proc., the World Congress on Intelligent Control and automation, 2804-2809

[5] Ishii, A., and Tomita, K. (2006) “Total Design of a Double Fronts Work Machine.” Proc., the 2006 JSME Conf. on Robotics and Mechatronics, 2A1-B07 (in Japanese)

[6] Ishii, A. (2006) “Operation System of a Double-Front Work Machine for Simultaneous Operation.” Proc., 23rd Int. Symp. on Automation and Robotics in Construction, 539-542

[7] Kamezaki, M., Iwata, H., and Sugano, S. (2008) “Development of an operation skill-training simulator for double-front work machine.” Proc., IEEE / ASME Int. Conf. on Advanced Intelligent Mechatronics, 170-175

[8] Laet, T. D., Schutter, J. D., and Bruyninckx, H. (2008) “Rigorously Bayesian Range Finder Sensor Model for Dynamic Environments.” Proc., IEEE Int. Conf. on Robotics and Automation, 2994-3001

[9] Murphy, K. (2002) “Dynamic Bayesian networks: representation, inference and learning,” PhD thesis, University of California, Berkeley

[10] Open Dynamics Engine, Russell, S. http://www.ode.org/

[11] Rabiner, L. R. (1989) “A tutorial on hidden Markov models and selected applications in speech recognition.” Proc., the IEEE, vol. 77, no. 2, 257-286

[12] Rodriguez, J., et al. (2004) “Operating experience of shovel drives for mining applications.” IEEE Trans. on Industrial Applications, 664-671

[13] Sakurai, Y., Nakada, T., and Tanaka, K. (2004) "Design method of intelligent oil-hydraulic system (load sensing oil-hydraulic system." Proc., IEEE / RSJ Int. Conf. on Intelligent Robot and System, 626-630
[14] Sasaki, R., Miyata, T., and Kawashima, K. (2004) "Development of remote control system of construction machinery using pneumatic robot arm." Proc., IEEE / RSJ Int. Conf. on Intelligent Robot and System, 748-753
[15] Sun, L. W., Meer, F. V., Bailly, Y., and Yeung, C. K. (2007) "Design and Development of a Da Vinci Surgical System simulator," Proc., the 2007 IEEE Int. Conf. on Mechatronics and Automation, 1050-1055
[16] Takemoto, A., Yano, K., Miyoshi, T., and Terashima, K. (2004) "Operation assist control system of rotary crane using proposed haptic joystick as man-machine interface." Proc., the IEEE Int. Workshop on Robot and Human Interactive Communication, 533-538
[17] Yoneda, M., Arai, F., Fukuda, T., Miyata, K., and Naito, T. (1997) "Operational assistance system for crane using interactive adaptation interface -design of 3D virtual crane simulator for operation training -." Proc., IEEE Int. Workshop on Robot and Human Communication, 224-229

Basic Technology toward Autonomous Hydraulic Excavator

Hiroshi Yamamoto[1], Masaharu Moteki[2], Hui Shao[3], Takashi Ootuki[4],
Humihiko Kanazawa[5] and Yoichi Tanaka[6]

[1234]Construction Technology Research Department, Advanced Technology Research Team, Incorporated Administrative Agency, Public Works Research Institute, 1-6, Minamihara, Tsukuba City, Ibaraki Pref., Japan; PH81-29-879-6757; FAX879-6799
[1] e-mail:h-yamamo@pwri.go.jp 2 e-mail:moteki@pwri.go.jp
[3] e-mail:shao77@pwri.go.jp 4 e-mail:ootuki@pwri.go.jp
[56] Research Center for Advanced Information Technology, Information Technology Div., National Institute for Land and Infrastructure Management, 1, Asahi, Tsukuba City, Ibaraki Pref., Japan; PH81-29-864-4916; FAX864-2690
[5]e-mail: kanazawa-f87bh@nilim.go.jp [6]e-mail: tanaka-y8317@nilim.go.jp

Abstract

Civil engineering construction work has always been accompanied by a high proportion of tasks that are either dangerous or unpleasant or both. Enhancing the general working environment and boosting safety levels are critical issues for the industry. Meanwhile, the industry has been slow to embrace IT, and there is substantial scope for the use of technology to boost efficiency, cut costs and improve quality levels in construction. In a bid to address this issue, the Ministry of Land, Infrastructure and Transport launched a five-year project in FY2003 entitled Development of Construction Robots and Associated IT Systems.

This paper reports on the research and development work carried out by the authors in connection with the project. The project developed a three-dimensional space description method for civil engineering execution, which can now be applied to finished work control etc. and used by construction machinery. Foundation technologies needed for self-controlled excavation and loading by a hydraulic excavator have been developed and verified to the level that, under basic conditions, they have attained speed and precision equal to that by a human operator.

Keywords: execution technology, information technology, robot technology, three-dimensional information, construction equipment

1. Outline of, and Background to, the Project

Civil engineering work is still performed in inferior work environments and involves dangerous and grueling tasks. And the introduction of IT is behind that in other industries, demanding the development of technologies to increase work efficiency, cut costs, and improve quality. Some researches on this field have produced a number of results, such as the research about the Automatic Excavator. [4] [7] [8]

The Ministry of Land, Infrastructure, Transport and Tourism carried out the General Technology Development Project, Development of Robots and other IT Execution Systems, as a five year plan extending from FY 2003 to FY 2007. This project applied the most advanced information technology (IT) and robot technology (RT) to carry out two projects, Development of IT Construction Machinery Execution Technologies and Development of Execution Control Technologies Using 3D Information, in order to eliminate dangerous and grueling tasks, improve inefficient work, and deal with similar challenges facing the field of civil engineering.

The goals of the research were to apply IT, RT etc. to earthwork in order to achieve execution control such as control and inspection of completed work based on three dimensional design and land form information, plus IT execution including supporting and automating the operation of construction machinery (see Fig. 1) [12]

This report presents an outline of the achievements of this project regarding the self-controlled operation of hydraulic shovels.

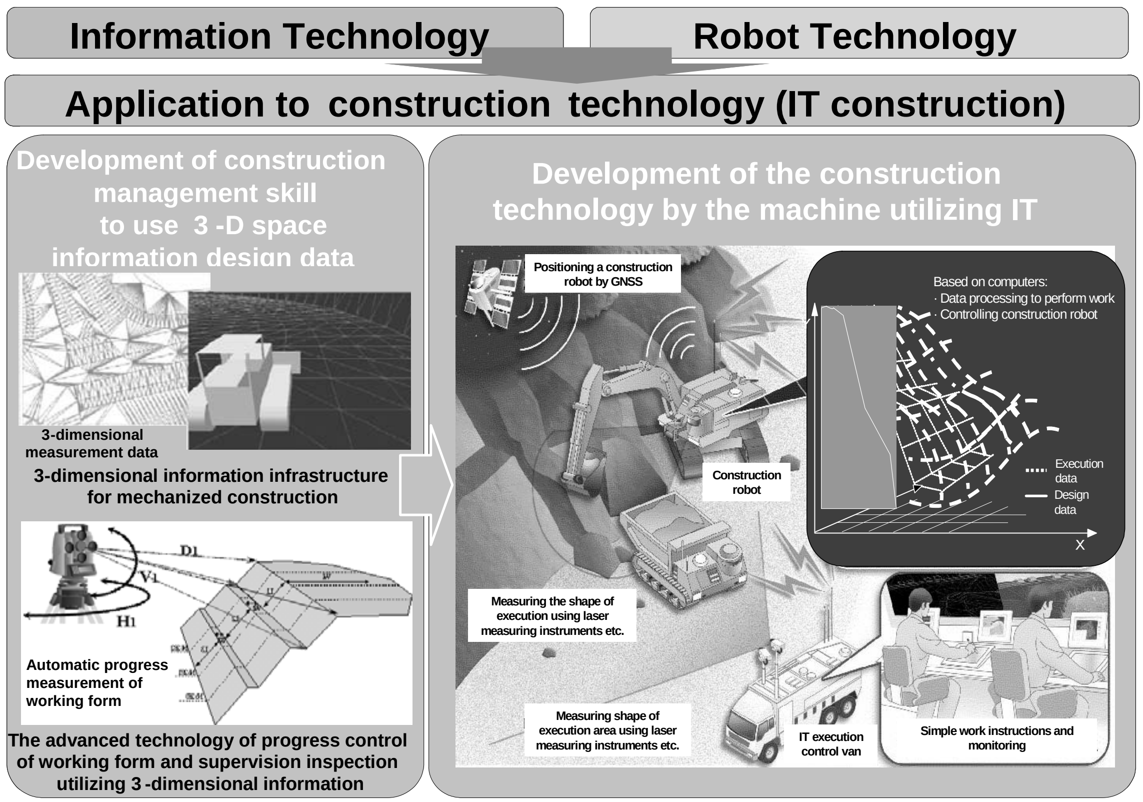


Figure.1 R/D of Advanced Execution Technology by RT and IT

The project was undertaken under the following policies.

1) Ministry of Land, Infrastructure, Transport and Tourism's Basic Technology Plan (2003 to 2007)
2) Integrated Science and Technology Policies: Next Generation Robot Related Policies (from July 2004)
3) Third Basic Science and Technology Plan (2006 to 2010)
4) Outline of the Promotion of Innovations in National Land and Transportation (May, 2007)
5) Information Integrated Construction Promotion Strategy (July 2008)

The Basic Science and Technology Plans are implemented to achieve the following research and development goals.

1) To apply three-dimensional design and land form information to develop IT execution systems for a robotic construction machine capable of automated excavation, and to improve construction site measurement and execution efficiency through remote operation by the end of 2007.
2) To apply automatic functions and measurement functions of construction machinery to improve safety and work productivity on construction sites and to achieve forms of execution capable of eliminating supplementary work by human workers by the end of 2010.

2. Development of Execution Control Technologies Using Three-dimensional Information

The goal is to perform more efficient execution control using three-dimensional design data and land form information obtainable in three dimensions as execution control information. The project has constructed execution control data which permits planar execution control to be installed in robotic construction machinery, prepared guidelines and handbooks for completed work control applying three-dimensional information [9], and has prepared written required development specifications necessary for the private sector to develop total stations (below called TS) [1].

2.1 Contents of the research which has been undertaken

(1) Clarification of the work to be improved and construction of an information model using three-dimensional design information

1) A proposal has been made to improve work between customers and contractors by IT—design verification, finished work control, supervision and inspections etc.—centered on three-dimensional design information which is a premise of an IT execution system.

2) An information model which shares both design and execution information between CAD and measuring equipment has been constructed.

3) Corroborative testing (4 locations) has verified the installation of finishing stakes and finished work control methods using three-dimensional information

4) Effectiveness of three-dimensional design information in executions—shortening the time required for preparatory work such as coordinate calculations etc.—has been confirmed.

(2) Improvement of information models using three-dimensional basic design information

1) Present three-dimensional measurement technologies (GPS, TS, three dimensional scanners, etc.) were studied, confirming the effectiveness of TS based finished work control.

2) Three-dimensional basic design information installed in measuring instruments has improved descriptive methods (information models) capable of responding to any changes of lateral section shapes from among design specifications or framework architecture such as widths or road alignments of structures

(3) Verification at construction sites

1) On-site trials of finished work control were performed using TS incorporated into basic design information (at six sites).

2) For the trials, three-dimensional design information was prepared, and at the site, this three-dimensional design information was compared with the finished work measurement results, verifying the series of procedures called automatic preparation of finished work control ledgers.

2.2 Research achievements

(1) Three-dimensional space data conversion specifications

By defining basic design information which consists of shape structure elements, basic design elements, and basic coordinate system elements, the research constructed three-dimensional design specifications which can be applied to TS-based finished work control at construction sites. (See Fig. 2) And based on the defined basic design information, three-dimensional design information needed by robotic construction machines was completed as three-dimensional space data conversion specifications (skeleton data). Three-dimensional space data conversion specification was expressed data elements based on land-XML used in the construction field. [2]

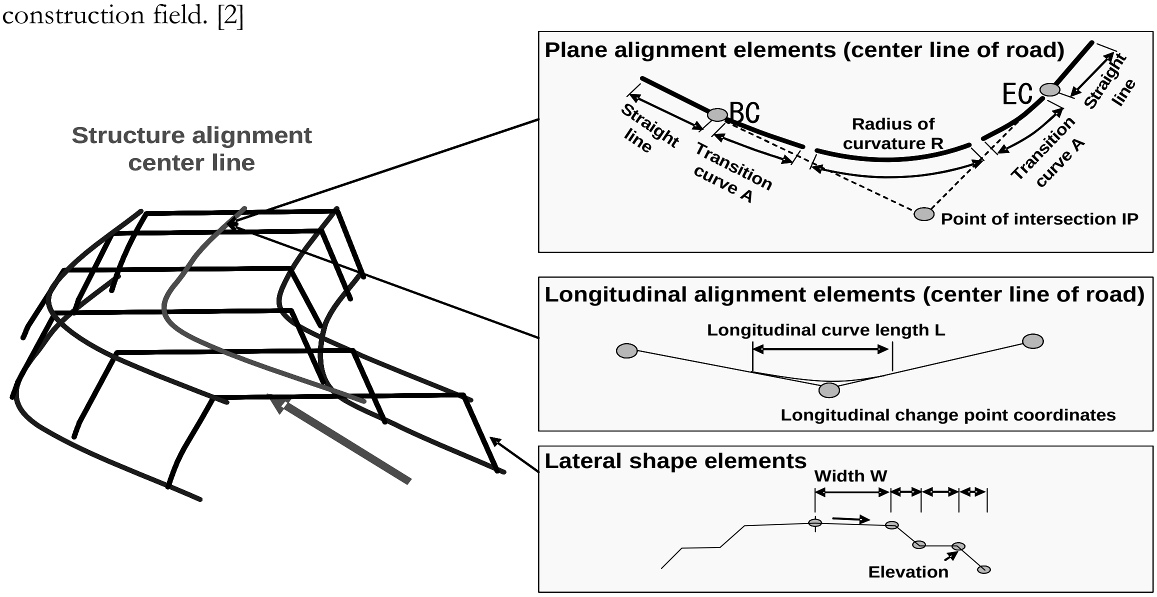


Figure 2. Three-dimensional Design Information

(2) Finished work control technologies and completion inspection technologies applying three-dimensional design information

Determining a new finished work measurement method based on TS incorporated in basic design information has completed the development of finished work control technologies and completion inspection technologies which apply three dimensional design information. (See Fig. 3) With this method, it was possible to confirm discrepancies between design values and measured values on site by calculating the slope length, width, length and height of berms similarly to the present method based on distance from and comparable height difference from the road center line. And it has also been verified that TS based measurements can be made more efficiently than measurements using levels or measuring tapes.

(3) Required specifications of equipment needed to handle three-dimensional information

Judging from results of the on-site trial, conditions necessary for TS to perform finished work control have been clarified as required hardware conditions and as required software conditions. And these have been summarized as specifications required by three-dimensional information application instruments in order that all measurement instrument makers will be able to develop equipment

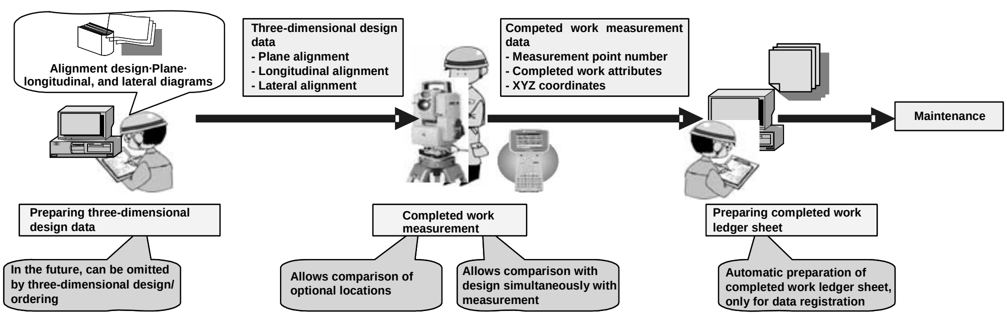


Figure 3. Finished Work Control Technology Using a Total Station

3. Development of Self-controlled Excavation and Loading Technologies for Hydraulic Excavators

3.1 Outline

To create practical IT execution technologies for construction machinery, the basic technologies which are its foundations will be developed. Specifically, the research will develop technologies to display three-dimensional design information and three-dimensional land form information as it changes according to the execution on an operation screen to provide work location, work contents, and other simple instruction information on a screen.

The goal is to introduce this as a work support system which improves execution efficiency of normal construction machinery and at the same time, develop IT execution technology based on robot construction machines such as hydraulic excavators which apply IT and robot technology to perform executions automatically.

The object of the research was to create foundation technologies in order to introduce IT and RT wherever possible within a practical range to earthwork at extremely dangerous work sites such as disaster restoration and disaster prevention projects (sediment control projects for example), as technologies necessary for execution processes to excavate, load, and transport soil using hydraulic excavators or crawlers which are representative general purpose construction machines. This was verified by a virtual site test as a prototype system. And the challenges revealed by the tests will be the grounds for future development.

The research and development is hypothesized as a two-stage process based on the maturity and practicality of the technologies.

1) The first stage goal is to develop a system to support work by presenting operators of construction machinery (remote control etc.) with the goal of the work (design) and present situation (land form) as three-dimensional information and with the position of the machine. This system is intended to be technology which will soon reach the practical stage at actual unmanned execution sites etc.
2) The second stage goal is to create a prototype of a robot construction machine which, when a remote operator provides simple work instructions including the location, range, and contents,

performs self-controlled work to a certain degree based on three-dimensional information concerning goals of the work (design) and present situation (land form).

"To a certain degree" in the above paragraph means that the operator intervenes to operate the machine remotely when it is difficult for the machine to perform self-controlled operation because, for example, an unexpected situation such as a large rock occurs. And self-controlled work means that the bucket, boom and other working parts of the hydraulic excavator plan the work and perform self-controlled excavation and loading work while revising the plan according to the situation.

The goal of the research and development is a system permitting a machine to, based on three dimensional design and shape information (three-dimensional space data conversion specifications), measure the land form of the ground as it is altered by the work, plan the excavation and loading work according to the measurements and to control the machine so it performs self-controlled work with approximately the same precision and at about the same speed as an experienced operator.

3.2 Contents of the research and development

To achieve this goal, three kinds of technology were developed: (1) technology which recognizes the surrounding environment as three-dimensional information necessary for self-controlled execution, (2) technology which displays three-dimensional information necessary for the work (man-machine interface), and (3) technology to automate execution (control technology).

This was done under limiting conditions: that soil and ground conditions be almost uniform at the research institute test site, the work contents be trench excavation and loading, the shape of the work be based on design information, the work range be designated, and the movement be remotely controlled. Figure 4 is a schematic diagram of the prototype which was developed, and Figure 5 shows an exterior view of the hydraulic excavator and the locations where sensors are installed.

The basic concept of this research and development is not accumulating the essence of advanced technologies even if at great cost, so if it achieves its goals, it is a success; rather it is to achieve the minimum level of target functions at the lowest possible cost by, instead of using special basic technologies, applying technologies which are as long-established as possible. Because it is a rapidly progressing field of technology, the functions were divided into sub-systems and components to construct it by simplifying the replacement of these with higher functioning or less expensive sub-systems and components in order to easily apply new technologies to each sub-system or component. [13] The development, which had a deadline for completion, was done by omitting model tests to use actual construction machinery from the beginning. The hydraulic excavator used as the base machine was, small during the testing period and its equipment installation capacity was large, so a 12t class machine was used. And even though it as a test machine, as safety measures, the testing was performed carefully using emergency stop devices and monitoring communications. [3]

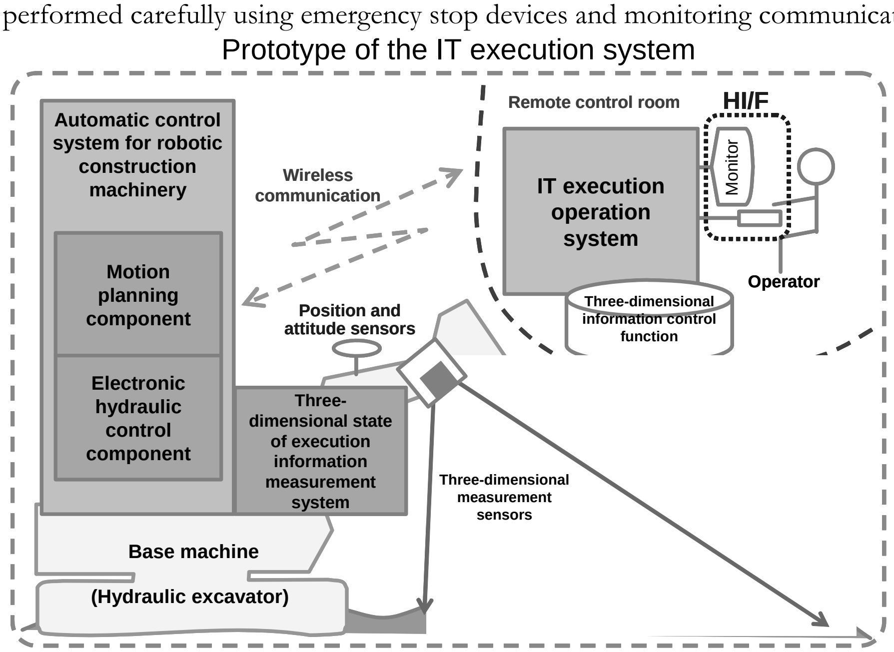


Figure 4. Schematic Diagram of the System

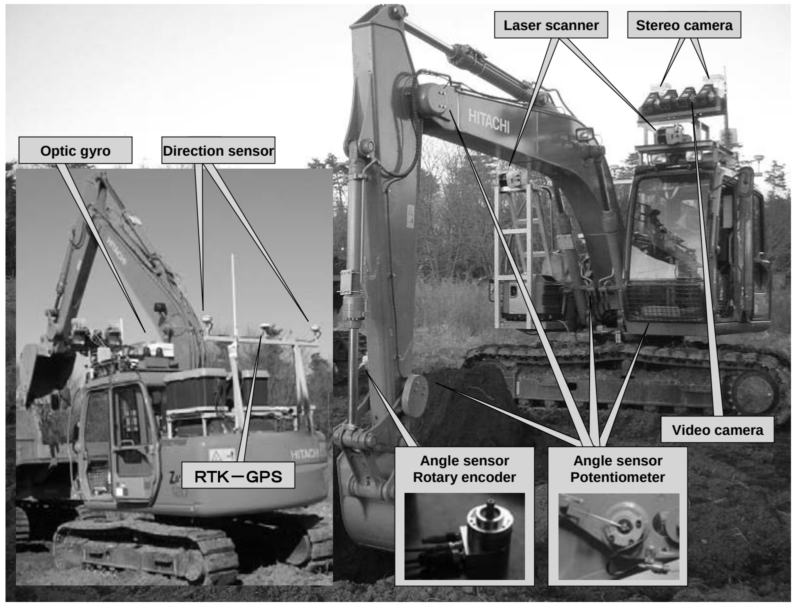


Figure 5. Hydraulic Excavator: External Appearance and Locations of Sensors

3.3 Outline of the prototype

(1) The three-dimensional state of execution information measurement system

The location of the machine is measured by RTK-GPS, its direction, inclination, and oscillation by optical fiber gyros, the directional gyro drift correction is done by a GPS phase difference direction meter. The accuracy of Optic gyro measuring attitude angle is ±(0.2°+1% of angle). And the ground is measured by a laser scanner (real time) and a stereo camera (finished work confirmation) etc. The maximum measuring range of the laser is 80m with 10mm measurement resolution. The measurement is done by performing time management and synchronization to convert the machine coordinate system and world coordinate system. [11]

(2) Display and operating systems

The results of measurements of machine position, design, and ground surface measurements are displayed by computer graphics to support autonomous setting, instructions, monitoring, and other remote operations. And the moving images obtained by the camera are superimposed on the design lines. [14]

(3) Three-dimensional information control system

Design data is stored by three-dimensional space data conversion specifications. The design and ground surface measurement results are maintained by converting them to simplified meshes (5 - 10cm) and providing them to the various components.

(4) Automatic control system

The internal sensors in the working devices (bucket, boom, arm) are a rotary encoder, potentiometer, and stroke sensor. The control performance achieved with the encoder that the resolution is 14bits and the linearity error is ±1/2·LSB. The automatic bucket movement plan generating algorithm [6] and hydraulic control technology were developed with reference to operator operations analysis results [5] [10].

4. Verification by the Prototype

A prototype test on a simulated site on the grounds of the institute verified the system's functions by using a hydraulic excavator and a crawler dump truck to perform mechanized earthwork: excavation and loading soil. (Fig.6, Fig.7) To perform accurate movements preventing collisions with the crawler dump truck, the cycle time was set at approximately 30 seconds for preliminary excavation and approximately 40 seconds for finishing excavation, achieving finishing excavation precision of 5cm, verifying and confirming it as a foundation technology. The challenges were achieving smooth movement, shortening cycle time, and dealing with the problem of soil which could not be scooped up spilling from the bucket

Figure 6. Excavation and Loading

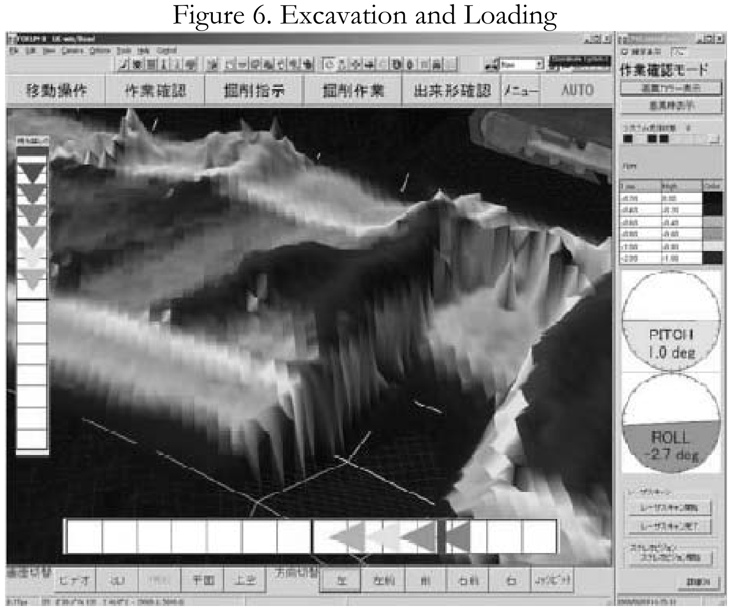


Figure 7. Display of Camera Measurements After Excavation

5. Conclusions

This project successfully established IT construction of earthwork—execution control including finished work control and inspections and support for and automation of the operation of construction machinery—using IT and RT and by applying three-dimensional design and land form information

Future endeavors in the area of execution control technologies using three-dimensional information will achieve finished work control technology using RTK-GPS by developing finished work control technologies for work categories other than earthwork and measuring instruments other than TS. It will be necessary to support information circulation using finished work control technologies for uses other than execution.

Future endeavors in the field of IT execution technology for construction machinery will permit its application to diverse soil conditions and work contents, simplifying expansion of its functions. Technologically, evaluating ground properties in real time to reflect the results in movement planning and control, and achieving methods of presenting easily expandable automatic movement planning creation rule description methods. And it will be necessary to clarify effective operation support, particularly execution methods, by making use of self-control functions and measurement functions.

Acknowledgments

The authors received assistance with this research from many members of organizations such as the Ministry of Land, Infrastructure, Transport and Tourism, the Research Committee for IT Execution Systems Based on Robots Etc. (Committee Chairman: Professor Yuta of Tsukuba University), Surveying Instruments Manufacturer's Association, Japan Construction Mechanization Association, Advanced Construction Technology Center, the Construction Robot Committee of the Japan Society of Civil Engineers, and the Unmanned Construction System Association and others. We are deeply grateful to them all for their support. And we also wish to thank everyone who assisted with the testing.

References

[1] Kambara, A., Tanaka Y., Kanazawa F. (2008),"Approach of as-built management by Total Station in earthworks" Proc., 11th Symposium on Construction Robotics in Japan, 173-182.

[2] LandXML.org: http://www.landxml.org

[3] Nozue A., Yamamoto H., Sakaida Y., Shao H., Yanagisawa Y. (2008) "About the application of the communication information system in the development of the IT construction technology of the hydraulic excavator -Profit use of network technology-" Proc., 11th Symposium on Construction Robotics in Japan, 277-284.

[4] Rowe, P., Stentz. A. (1997). "Parameterized scripts for motion planning." Proc. International Conference on Intelligent Robots and Systems,1119– 1124.

[5] Sakaida Y., Yamamoto H., Shao H., Nozue A. , Yanagisawa Y. (2008) "Analysis of Skillful Hydraulic Excavator Operation-Individual variation analysis between operators–", Proc., 11th Symposium on Construction Robotics in Japan, 263-270.

[6] Shao H., Yamamoto H., Sakaida Y., Yanagisawa Y., Nozue A. , Yamaguchi T., "Research on Autonomous Excavation and Loading Planning of Hydraulic Excavator" Proc., 11th Symposium on Construction Robotics in Japan, , 271-276.

[7] Singh S. (1995). "Synthesis of Tactical Plans for Robotic Excavation." doctoral dissertation, Robotics Institute, Carnegie Mellon University.

[8] Stentz A., Bares J., Singh S., and Rowe P. (1999). "A robotic excavator for autonomous truck loading." Autonomous robots, Kluwer Academic Publishers, Hingham, MA, USA, Volume 7 , Issue 2, 175 - 186.

[9] Tanaka Y., Kanbara A, Kanazawa F. (2008) "Development of construction management data using on the construction site" Proc., 11th Symposium on Construction Robotics in Japan, 163-172.

[10] Yamaguchi T., Yamamoto H. (2006),"Motion Analysis of Hydraulic Excavator in Excavating and Loading Work for Autonomous Control" 23rd ISARC2006, 602-607.

[11] Yamamoto H., Ishimatsu Y., Ageishi,S., Ikeda,N., Endo,K., Masuda,M., Uchida,M., Yamaguchi,H.(2006),"Example of Experimental Use of 3D Measurement System for Construction Robot Based on Component Design Concept" 23rd ISARC2006, 252-257.

[12] Yamamoto, H., Ishimatsu Y., Yamaguchi T., Uesaka K., Aritomi K., Tananka Y. (2006), "Introduction to The General Technology Development Project:Research and Development of Advanced Execution Technology by Remote Control Robot and Information Technology" 23rd ISARC2006, 24-29.

[13] Yamamoto H., Yanagisawa Y., Shao H., Sakaida Y., Nozue A., Yamaguchi T. (2008), "Basic Technology toward Autonomous Hydraulic Excavator -Development of Construction Robots and Associated IT Systems-" Proc., 11th Symposium on Construction Robotics in Japan, 243-252.

[14] Yanagisawa Y., Yamamoto H., Shao H., Sakaida Y., Nozue A., Yamaguchi T. (2008). "Research on a Man-machine Interface for Remote Operation of Work Machinery" Proc., 11th Symposium on Construction Robotics in Japan, 253-262.

Automation in the Prefab and Modular Construction Industry

Joseph Neelamkavil

National Research Council Canada, 800 Collip Circle, London, Ontario, Canada N6G 4X8, Email: joseph.neelamkavil@nrc.gc.ca

Abstract

The construction industry at-large is dominated by numerous small and specialized sub-contractors who typically are not technologically advanced enough to embrace automation. The sector that represents factory built housing (the modular, prefab, panelized, precast, etc.) is an exception. Since the products are built in factories, the principles of mass production and mass customization that are the norm in manufacturing, apply. It will also make it easy to adapt to automation, integration and optimization. In this scenario, the constructability aspects can be verified prior to being built, and an optimum construction plan can be derived. Newer materials can be applied, tight tolerances achieved, while the built products are not affected by outside climatic conditions, as is the case for site built housing. With a view to gauge the needs of automation within the prefab sector, we reviewed several automation technologies relevant to construction, and also have been consulting with the prefab industry to understand their state of practice. This paper summarizes the findings of this study, provides an overview of the type of automation that is prevalent in the industry, and outlines a technology outlook with various elements comprising the aspects of prefab automation. It will help the prefab industry personnel arrive at a broad view of the essence of construction automation, while the research community gets a sense of industry 'pull' in embracing automation.

Introduction

The uniqueness of the construction sector poses several challenges for the direct adaptation of technologies that are used in many other industries, for example, those that support mass production and/or mass customization. This provides opportunities for researchers engaged in construction-related research. It also involves extensive use of information and support technologies across the enterprise and its market. The automation pertaining to the prefab construction sector, typically falls into one of three categories:1) prefab components making process, (parts, panels, precast, formwork, etc.) which deal primarily with the construction of the building blocks; 2) assembly process in which the construction components (often from different suppliers) are installed to create buildings, houses, etc. by an array of sub-contractors, sometimes having conflicting workflows; 3) construction business processes that represent both the business and support processes (project management, supply chain management, document management, workflow management, change management, planning & scheduling, etc.).

The FutureHome Project

A milestone European prefab automation project -The FutureHome - started in 1998 and was completed in 2002.The main objective of the FutureHome project represents the development of an integrated construction automation concept and associated technologies for all stages of the house-building construction process, from the architect's desk to site robots. This includes: a) the modular design of buildings with planned robotic erection, b) automatic planning and real-time planning of offsite prefabrication, transportation, and onsite assembly and c) onsite automated transportation, manipulation, and assembly of the prefabricated parts. The FutureHome building system derives from a Kit-of-Parts approach, which is a specific implementation of prefabrication. As noted by Wing et al (2002), virtual reality is employed, providing a common virtual environment that can be shared by clients, architects, engineers, constructors, etc. The clients can be led through the virtual design of their house, created on-line using "prefabricated" components. The construction process can be simulated, allowing examination of cost, quality and time aspects. This is an environment within which the user creates a design from its constituent prefabricated components. A 3D window lists a library of available components, which can be assembled automatically. The user selects a component and views it in 3D and chooses to add the component to the design, using collision detection and constraint-based modeling techniques to ease the interaction process. As noted in Dietz et al (2007), several assembly connectors had been developed for the assembly of the modules, the structural connection, and electrical and service pipes connections. These connectors ensure automatic performance of complete assembly between modules. A noteworthy feature that evolved as part of this project is the development of a tool called AUTMOD3 - an automatic modular construction software environment. This system integrates architectural design, planning and simulation tools in a commercial CAD program.

The ManuBuild Project

Another prefab automation project - The ManuBuild - started in 2005, is due to be completed in 2009. This European project, as referred in Bock (2006) and the Manubuild website (refer 31), has set its objective such that customers can purchase high quality manufactured buildings that have a high degree of design flexibility at relatively low cost. Enabling business processes, ICT systems, new materials and smart components are major features of this project. ManuBuild has set its vision: 'Customers will be actively engaged in the design of their buildings, using state of the art interactive tools; it incorporates mass customization, and offers customers increased choice and design flexibility; the efficient, flexible and scalable manufacturing concept enables production efficiencies of production scale; an open system for products and components gives diversity of supply at competitive costs'. The potential impact includes significant reductions in waste, costs, time to construct and the number of construction related accidents. Striving to achieve automation, ManuBuild aims to provide ICT support for distributed building manufacturing. To achieve this, several decision support tools have been developed, ranging from catalogues of products via information delivery, design, custom configuration, and assembly planning.

Other Major Projects:

There have been other major project undertakings also pertaining to construction automation, especially in the United States. For example, FIATECH (2004) has created a Capital Projects Technology roadmap (refer to Figure 1) to integrate various functions and any required information in a unified construction project/facility management environment, as briefed in Bowden et al. (2006). In our study, we limit the scope to the elements of automation that are a) Automated Design, b) Integrated Automated Procurement and Supply Network and c) Intelligent and Automated Construction Job Site.

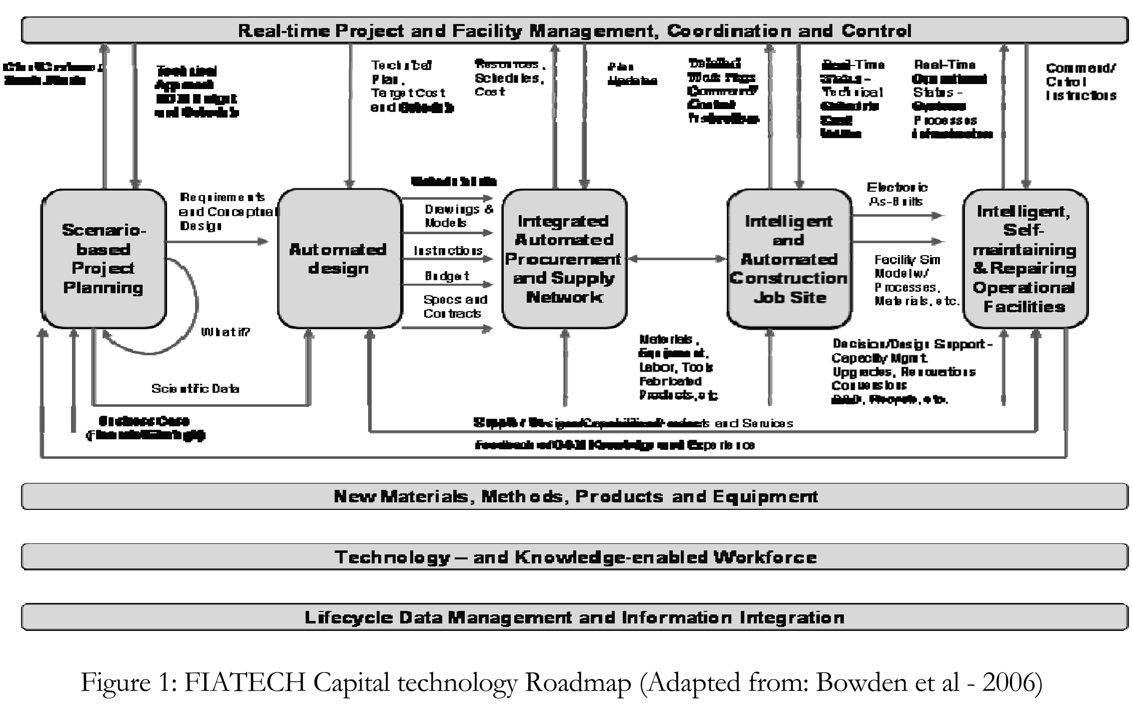


Figure 1: FIATECH Capital technology Roadmap (Adapted from: Bowden et al - 2006)

Automated Design

It may be argued that the construction starts at design phase of the built environment. The prime considerations of this phase is the advancement of automated design to improve cost-efficiency, enhance the lifecycle value of the project, and enable interoperability among a project's lifecycle entities, and all these can result in significant cost savings. Accordingly, technology integration, with capabilities in 3-D design, analytical modeling and simulation and distributed intelligence, offers a great opportunity to create a truly integrated and automated design environment. It will greatly reduce errors through automated design optimization and verification. Optimization could include a variety of options including installation cost, lifecycle cost, and plant output. It should process design options in an accurate, scenario-based visualization environment.

The present day 3D modeling tools embed pre-defined objects that facilitate the development, routing, and connection of building systems in 3D, and provide conflict detection to identify physical interferences between components. 4D modeling tools (3D CAD model with time as its 4^{th} dimension) link a construction project's scope in 3D with its schedule to simulate the actual construction process. The benefits include the elimination of construction interferences, less reworks, fewer change orders, increased productivity and a decrease in delivery time, as detailed by Staub-French and Khanzod (2007). By reviewing project schedules in a 4D environment, and as referred in Yerrapathruni et al (2003), the construction professionals can readily identify design, constructability, sequencing and interdisciplinary interfacing issues.

BIM and Data Exchange Models

Automated data exchange and wide usage of Building Information Model (BIM) play major roles in achieving construction automation. Technology wise, as noted in the Autodesk website (refer 29), BIM is an approach to building design that is characterized by the creation and use of coordinated, internally consistent computable information about a construction project. And, according to National Building Information Model Standard published by The National Institute of Building Sciences (NIBS), BIM is defined as "a digital representation of physical and functional characteristics of a facility. As such it serves as a shared knowledge resource for information about a facility, forming a reliable basis for decisions during its lifecycle from inception onward." The foundation of BIM is parametric building modeling, which records, presents, and manages not only object data at the component level, but also the network of relationships among all of the objects of the building from various views. From a lifecycle point of view, BIM enables architects, engineers, contractors, owners, and facility managers to share data throughout the entire lifecycle of the building. The shared data includes the initial design data; geospatial information; financial and legal data; mechanical, electrical, and plumbing layouts; building product specifications, environmental and energy modeling results, and other information, as described in McGraw_Hill Construction 2007 (refer 33). With BIM in place, large amounts of data (typical in moderate and large projects) can be continuously exchanged among the key players. This enables facts, figures, designs and analyses that affect one or more information sources to be constantly updated to ensure that any decisions made are based on accurate information.

Automated Supply Network and Materials Management

As detailed in Ganeshan and Harrison (1995), a supply chain is a network of facilities and distribution options that performs the functions of procurement of materials, transformation of these materials into intermediate and finished products, and the distribution of these finished products to customers. By properly managing the network, one ensures the supply of the right products in the right quantities at the right places at the right time and at a minimal cost. An important aspect of this setup is the 'coordination' of organizations. Xue et al. (2005) presented an agent-based framework for supply chain coordination, based on multi-attribute negotiation and utility theory. A notable feature here is the coordinated decision making process of multiple organizations through the integration of key business processes and key members involved in the construction supply chain, including client/owner, architect, general contractors, suppliers, etc. Complementary to the supply chain management is the materials management function. Navon and Berkovich (2005) have suggested a material management and control model that encompasses the purchase of materials, delivery to the construction site, and dispatching of materials for usage. Intelligent and automated data collection technology has been incorporated into this system for tracking materials, via enabling technologies like the use of barcodes, RFID, and/or personal digital assistants.

Robotics Automation in Prefab Industry

Robots found its way to construction first in components making and the production of modular housing. Later, mobile robots were developed for special on-site construction tasks. As reported by Bock (2004), robots played active roles at the production line of Sekisui Chemical Sekisui Heim in Japan, where more than 85 percent of the houses are prefabricated. Bock (2007) elaborates a robotic precast concrete panel factory that uses a multipurpose unit which allows flexible production of the concrete floor, wall and roof panels. Here, according to certain CAD data, a multi functional gantry type robotic unit with two vertical arms places magnetos on the steel production table. The unit also attaches shutters on top of the magnetos and then places horizontal, vertical and triangular reinforcement bars, as per design. A CAD-CAM controlled concrete distributor spreads the right amount of concrete while controlled by a CAD layout plan, which takes into account installation, window or door openings.

Kye-Young Lee et al (2006) studied the human-robot cooperative (HRC) system to cope with the construction environment via real-time interaction with human and robot simultaneously. The power of a robot system helps a human to handle heavy materials with a scaled-down load. The human can feel and respond to the force reflected from the robot end-effecter acting with the working environment. The study also presents the experiments and evaluations to verify the HRC system for robotic applications giving the performance tests, like force assistance, force reflection, and position-tracking in a two axes manipulator. Engelbert Westkamper et al (2000) developed a robotic system for the automatic laying of tiles within certain tolerances on prefabricated modules. The pilot work consisted of a tile laying system that consists of tile positioning equipment, a centering and measuring system and laying unit; an adhesive application system, consisting of adhesive preparation and transport unit; a tile supply system consisting of a store and a measuring unit; systems for generating process parameters; and handling and positioning systems having industrial robot and process control. Peñin et al (1998) and Paster et al (2001) developed a robotized cell for making the pre-fabricated glass reinforced cement panels for a Spanish construction company. It is developed for the automatic programming and control of facade panel manufacturing, and a CAD based 3D-drawing of the building facade serves as the input.

Automated Construction Site

An automated construction site may use robotics for logistics and assembly, but can face many barriers that are technological or economical. The technological barriers are that a robot must cope with the complexity of the construction process involving a dynamic and evolving site, together with the need for performing multiple tasks with differing characteristics. As detailed by Zied (2007), robotics research in construction has moved to address technological barriers via: a) development of mobile platforms and manipulators, b) development of control systems and sensory systems integration, c) re-engineering of processes to suit robotic systems, d) software development related to support the above, and use of advanced IT systems to enhance the whole system performance. Economical barriers also influence the implementation of a robotic system in construction, such as: a) the cost vs. benefits, b) the changes required for implementing the new system, and c) the effect of the new system on the complete organization, which includes health and safety, and labor unions concerns.

Martinez et al (2008) have reported on the robotization and industrialization of the construction process, with objectives on: a) Modular design of buildings, with robotic erection in construction; b) automatic planning and real-time re-planning of the offsite prefabrication, transportation, and onsite assembly; c) onsite automated transportation and assembly of prefabricated components. They also have developed tools for the assembly of various building components and modules. SMART (Shimizu Manufacturing system by Advanced Robotics Technology) represents a systems approach (reference 34) to computer-integrated construction. The system set-up takes a few weeks, after which the building's top floor and roof are erected on top of four jacking towers. Jacking towers are used to push up the several-ton heavy top floor assembly - the main work platform - as well as lifting their own bases from floor to floor. The system is also composed of lifting mechanisms and automatic conveying equipment which are installed on the work platform. Overhead gantry cranes are connected to the underside of the roof structure. Trolley hoists are used to lift and position prefab components introduced at ground level, and floors emerge from under the top floor. The whole process is computer-controlled, though site workers oversee the operations.

Wakisaka et al (2000) have described a similar construction system for high-rise reinforced concrete buildings, which is often referred to as the 'Big Canopy'. Its set up consists of a parallel delivery system with three automated overhead cranes and one large construction lift under a synchronous, all-weather, climbing, temporary roof frame. A material management system is also part of it, with database linkage and communication with a CAD system, which draw resources based on prefabricated components. The Big Canopy construction system is divided into the following subsystems: a) a roof supported by four tower crane posts that are situated outside the building; b) a complex hoist system with three cranes mounted against the roof; c) a jib crane on the roof to mount and to dismantle the tower crane posts; d) a high-speed lift to all floors; e) use of prefabricated components with easy identification (RFID, etc.); and f) a material and delivery system to manage the flow of materials and components.

Virtual Reality and Simulation in Prefab Industry

Virtual reality (VR) enables real-time viewing of, and interaction with, spatial information. The visualization can simulate performing such tasks as production, transportation, handling and assembly of components. It can also evaluate which pieces of the building are to be assembled in what sequence, as well as reuse the model at different stages of construction. Team members can visualize their specific tasks and the relationships between the works of various other sub-contractors. Li et al (2008) described an integrated framework and process for general contractors to apply the virtual prototyping technology. In this framework, an expert (process modeler) accepts the BIM model from the designer, and decomposes it into formats required by contractors and consultants. At the same time, the process modeler integrates information provided by a team member into the BIM model to create a virtual prototyping of the construction processes. Through an iterative process, the process modeler enables the construction team to conduct 'what-if' analyses of different construction methods in the virtual prototyping environment until a satisfactory method is achieved.

Murray et al (2003) described a virtual environment that supports the design, and the assembly of a building from prefabricated components. The environment imports a library of 3-D models of prefabricated modules to interactively design a building. The construction schedule can be altered, and the information may be fed back to its simulation environment. The users can create their design through automatic constraint recognition and view the real-time construction process simulation. The site manager viewing the system gets a good project perception, as it provides the project's schedule management function integrated with an animated virtual reality display. Nasereddin et al (2007) described an automated approach to create a discrete event simulation model, developed using ProModel and Visual Basic. It is an approach for automating the model development for the prefab industry. A case study is presented to verify advantages that include a reduction in model development time and improved modeling consistency. Here, the animation data which was input manually includes the graphics library that they had created for the modular home industry. The library is used to show animation changes to the module as it is processed through the production line.

Planning for Prefab Automated Assembly

A building is comprised of components – the basic building blocks - assembled as per certain rules and constraints. And in some sense, the prefabricated building construction is analogous to the assembly of a manufactured product; if we consider assembly as the reverse of disassembly, the whole assembly process sequences may be inferred. But, the assembly of components is dependent not only on geometric properties, but also on the assembly relationship and spatial restriction information of the components. Hu (2005) has developed a method for the automatic determination of the construction sequence from a connection graph representing a pre-fabricated building via the extraction of components, and by a verification of disassembly. Components tightly connected geometrically and physically, but loosely connected to the rest of the assembly, are extracted as a preferred super-component. The process of extracting components not only makes it possible to reduce the problem space by early pruning of infeasible construction sequences, but also explicitly defines temporal and spatial parallelism in construction. Benjaoran and Dawood (2006) developed an innovative planning system called 'Artificial Intelligence Planner' (AIP) that helps automating the planning segment of 'bespoke' precast production. Its data integration features encourages automation in the planning process, and its decision support capability enables planners to improve the production plan efficiency. These two complement each other to deliver optimum benefits to precast manufacturers. The AIP system adopts artificial intelligence technologies (neural network and genetic algorithm) to assist the process of production planning.

Scheduling Automation and Sensor-based Control

An efficient production schedule drives automation in the actual creation of prefab components. Several scheduling prototypes have been reported, which addresses this issue. For example, a decision support system for coordinated prefabrication scheduling has been described by Chan and Zeng (2003). It supports key elements of production (re)scheduling, namely, conflict detection, determination of the priority for conflict resolution, generation and evaluation of alternatives for conflict resolution, and ranking of outcomes for negotiation. It combines the use of an explicit constraints-based scheduling model, and genetic algorithms (GA) to determine scheduling parameters and conflict resolution priorities. A GA based searching technique is also adopted in a mixed precast flow-shop scheduling system proposed by Leu and Hwang (2002), providing near-optimal combination of production sequences, resource utilization, and minimum make-span while complying with resource constraints and mixed production. Take notice that prior to construction, the components delivered to the site need to be inspected, stored and then tracked, while the details related to installation require adequate documentation. Prefab components locations need to be known, and transported to the site with minimum time loss. Gajamani and Varghese (2007) have presented a RFID-based location tracking real time scheduling and monitoring system that make use of prefab/ precast components. Data is collected in real time and is converted using a software system, not only for the control but also for field material management and future use.

Conclusions & Future Directions

The ultimate objective of automation is to improve productivity, quality and safety, which will also contribute to cost reduction. The type of automation technologies detailed in this study revolves around actual construction projects as well as research activities that are related to building systems: that is, construction components, fabricated modules, panels, houses etc. As is somewhat obvious, the three prime areas of automation in the prefab and modular construction domain are: the design, the material handling (robots, etc.) and the business processes (including planning and scheduling). It is logical to conclude that some form of automation will be the norm in all factory-built housing of the future. Research directions in the form of a roadmap for the prefab industry is sketched out below, based on this study, and our interactions with the regional construction industry, as well as our own understanding of the type of automation that prevails in other sectors like manufacturing. This conclusion has been reinforced by the findings given in Mullens (2004), the technology roadmap by the manufactured housing research alliance (refer 32), and the summary based on a consultant's report (refer 30). The roadmap details primarily with the productivity issues of factory settings (where the components are built), and the quality and functionality of the built components and their subsequent assembly that constitute the buildings.

Prefab Factory Automation – A Productivity Issue

Embracing the technologies that revolutionized manufacturing industry, prefab builders may radically improve the efficiency with which it produces end products - buildings. The future builders will evaluate ways to improve productivity through the application of lean production, information and automation technologies. It will include efficient methods for warehousing, develop strategies to reduce construction waste, and adopt techniques for recycling. It will develop and deploy technologies for defect-free transportation of the prefab components. It will develop and test transportation system features that are high performance, low cost and more fully integrated into the building's structural system. Support processes will include optimization, simulation, visualization and project management tools to evaluate and control alternatives. Materials will arrive in the factory just in time to support production, and stored close to the point of use. Components will be identified and tracked in real time to monitor the exact location and status of all items. Material handling and manufacturing processes will be automated to eliminate injuries, minimize physical exertion, assure capacity and boost productivity. Production documentation will be accurate and timely, and can be accessed from remote locations. Employees will know the status of any order, recognize the limitations of any resources, and will be able to react to meet schedule commitments. Product rework will be minimal and production flow will be smooth and synchronous with demand.

Prefab Construction Automation – A Quality and Functionality Issue

Technological advances to improve the quality of a product in terms of its function, is critical. Forward thinking on how the prefab product of the future will look and behave gives good insight. Prefab modules and buildings will adopt a systems integration approach to design, engineering and construction. The building envelope systems (roof, wall, floor, etc.) will maximize integrated performance. Builders will incorporate fully integrated structural systems that can resist the forces applied during component/module transport and site installation. It will develop new assemblies and sub-assemblies that improve performance that could be applied with ease in a factory setting. It will integrate plumbing and mechanical systems within the whole module, and will include new wiring and cabling systems that optimize whole-building performance. It will also incorporate functions that wired and wireless technologies and systems can offer. It will identify the reasons for material failures and/or systems linking to manufacturing and/or installation problems. It will develop new production, installation and joining (ie: fasteners, joints, etc.) techniques, and new designs to minimize failures. A virtual and simulation environment that supports the design of prefab components and their assembly, and simulating the construction to suit schedule will become the standard practice. The simulation environment likely will utilize a library of 3-D models of the prefab modules that can be used to interactively design the whole building. The design and operation of the building will promote and contribute to the health of its occupants. Buildings will have deployed methods for controlling sources of contamination, techniques for improving ventilation, and systems & procedures for controlling moisture. It will develop and deploy new and existing energy conservation and renewable energy technologies and strategies. The built environments will be the most energy efficient, and annual energy costs of these units will be low, or perhaps even lower than comparable site-built structures.

References

[1] Benjaoran, V.; Dawood, N (2006) "Intelligent approach to prod. planning system for bespoke precast concrete products", Automation in Construction 15, pp 737 – 745.

[2] Bock, T. (2004) "Construction robotics and automation: past-present-future", World Automation Congress, Proceedings June 28 2004-July 1, 2004.

[3] Bock, T (2006), Tech. University Munich, http://www.iaarc.org/external/isarc2006-cd/www/ISARC2006/00174_200606072132.pdf

[4] Bock, T (2007) "Construction robotics", Auton Robot (2007) 22:201–209

[5] Bowden, S., Dorr, A., Thorpe, T., and Anumba, C. (2006). "Mobile ICT support for construction process improvement." *Automation in Construction*, 15, 664–676.

[6] Chan, W.T.; Zeng, Z. (2003) "Coordinated Production Scheduling of Prefabricated Building Components", Construction Research, Winds of Change: Integration and Innovation of Construction.

[7] Diez, R, Padrón, V. M., Abderrahim, M, Balaguer, C (2007) "AUTMOD3: The Integration of Design and Planning Tools for Automatic Modular Construction", International Journal of Advanced Robotic Systems, Vol. 4, No. 4.

[8] Gajamani, G.K.; Varghese, K. (2007) "Automated Project Schedule and Inventory Monitoring using RFID", 24th International Symposium on Automation in Robotics in Construction – ISARC 2007.

[9] Ganeshan R., and Harrison T. P. (1995), Penn State University "An Introduction to Supply Chain Management".http://lcm.csa.iisc.ernet.in/scm/supply_chain_intro.html.

[10] Hu, W (2005) "Automatic Construction Process of Prefabricated Buildings on Geometric Reasoning", Proc: Construction Research Congress, San Diego USA.

[11] Lee, K.Y; Lee, S.Y; Choi, J.H; Lee, S.H; Han, C.S. (2006) "The application of the human-robot cooperative system for construction robot manipulating and installing heavy materials", SICE-ICASE International Joint Conference, Busan, Korea.

[12] Leu, S.S.; Hwang, S.T. (2002) "GA-based resource-constrained flow-shop model for mixed precast production", Automation in Construction 11, pp 439– 452

[13] Li, H.; Huang, T.; C.W. Kong, H.L. Guo; Baldwin, A.; Chan, N.; Wong, J. (2008); "Integrating design and construction through virtual prototyping", Automation in Construction 17, pp 915–922.

[14] Martinez, S; Jardon, A.; Navarro, J.M; Gonzalez, P. (2008) "Building industrialization: robotized assembly of modular products", Assembly Automation, Vol. 28, No. 2, pp 134–142.
[15] Mullens, M.A. (2004) "Production Flow and Shop Floor Control: Structuring the Modular Factory for Custom Homebuilding", www.pathnet.org/si.asp?id=1064
[16] Murray, N.; Fernando, T.; Aouad, G. (2003) "A Virtual Environment for the Design and Simulated Construction of Prefabricated Buildings", Virtual Reality 6:2 44–256, Springer-Verlag London Ltd.
[17] Nasereddin, M; Mullens, M.A.; Cope, D. (2007) "Automated simulator development: A strategy for modeling modular housing", Automation in Construction 16, 212–223
[18] Navon R., and Berkovich O. (2005) "Development and on-site evaluation of an auto. material management and control model", Journal of Const. Engg and Management.
[19] Pastor, J. M.; Balaguer, C; Rodriguez, F.J.; Diez, R. (2001), "Computer-Aided Architectural Design Oriented to Robotized Facade Panels Manufacturing", Computer-Aided Civil and Infrastructure Engineering, Vol 16, No. 3, pp. 216-227
[20] Peñin, L.F.; Balaguer, C.; Pastor, J.M. .; Rodríguez, F.J.; Barrientos, A.; Aracil, R. (1998) "robotized spraying cell of prefabricated panels in Construction industry", IEEE Robotics & Automation Magazine, Vol. 5. No. 3. pp.18-29.
[21] Staub-French, S., and Khanzod, A. (2007). "3D and 4D Modeling for Design and Construction Coordination: Issues and Lessons Learned", ITcon Vol. 12
[22] Wakisaka, T.; Furuya, N.; Inoue, Y.; Shiokawa, T. (2000) "Auto construction system for high-rise reinforced concrete buildings", Automation in Construction 9, 229–250.
[23] Westkamper E; Schraft, D.D; Laier, P. (2000) "Robot System for Automated Tiling on Prefabricated construction Elements", 31st Symposium on Robotics (ISR 2000).
[24] Whyte, J (2003) "industrial applications of virtual reality in architecture and construction", http://www.itcon.org/2003/4.
[25] Wing, R, Atkin, B. (2002) "FutureHome: A Prototype for Factory Housing", International Symposium on Automation and Robotics in Construction, 173-178 pp.
[26] Xue, X., Li, X., Shen, Q., and Wang, Y. (2005) "An agent-based framework for supply chain coordination construction", Automation in Construction 14, 413– 430.
[27] Yerrapathruni, S., Messner, J I., Baratta AJ., and Horman MJ. (2003) "Using 4D CAD and Immersive Virtual Environments to Improve Construction Planning" Construction and Virtual Reality conference CONVR 2003, USA.
[28] Zied, K. (2007) "An Augmented Framework for Practical Development of Construction Robots", International Journal of Advanced Robotic Systems, Vol. 4, No. 4, ISSN 1729‐8806, pp. 477‐500.
[29] http://images.autodesk.com/adsk/files/bim_parametric_building_modeling_jan07_1_.pdf - AutoDesk, Parametric building modeling: BIM's foundation, (2007)
[30] Consultant Report (restricted) - Altus Clayton (2008) "The Use of Prefabrication and Modularization in the Construction Industry".
[31] http://www.manubuild.org/downloads (2007)
[32] Manufactured Housing Research Alliance, (2003) "Technology Road Mapping for Manufactured Housing", March 2003.
[33] McGraw_Hill Construction (2007), Interoperability in the construction industry, SmartMarket Report, Interoperability Issue.
[34] "Quick Guide to Construction automation and robotics", http://adaptablehouse.vtt.fi/files/technology/robotics.doc.

Development of a Heuristics-based Task Planning System for Intelligent Excavating System

Seungsoo Lee[1], Jeonghwan Kim[2], Jinwoong Park[3], Jongwon Seo[4] and Youngsuk Kim[5]

[1]Research assistant, Dept. of Civil Engineering, Hanyang University, Seoul, Korea rokhonor99@gmail.com
[2]Research assistant, Dept. of Civil Engineering Hanyang University, Seoul, Korea, kimjh0418@gmail.com
[3]Research assistant, Dept. of Civil Engineering, Hanyang University, Seoul, Korea, okboard@hanyang.ac.kr
[4]Associate professor, Dept. of Civil Engineering, Hanyang University, Seoul, Korea, jseo@hanyang.ac.kr
[5]Associate professor, Dept. of Architecture, Inha University, Incheon, Korea, youngsuk@inha.ac.kr

Abstract

A research consortium called 'Intelligent Excavation System (IES)' has been formed in Korea with the support of the Ministry of Land, Transport and Maritime Affairs of Korea. The final goal of this research consortium is to develop a robotic excavator that can improve the productivity, the quality, and the safety of the conventional earthwork. The knowledge of the construction metrology and task level planning should be well fed into the robotic control mechanism in order for the machine to have the intelligence of the construction planner and the operator. Task Planning System (TPS) is one of the core technologies of IES. TPS generates an optimal earthwork system based on a virtual work environment updated in real time by work environment cognitive technology. TPS is an integrated module based on the heuristics of skillful excavator operators for effective and error-minimizing work planning. In this paper, the heuristics and the functions of system modules along with the virtual reality-based simulation results are presented.

Keywords: Intelligent Excavating System (IES), Task Planning System (TPS), Heuristics, Simulation, Earthwork

1. Introduction

1.1 Background and Purpose

The production lines in most industries now have been automated, which has brought many benefits such as the improvement of productivity and economical efficiency, the safety against industrial disasters and the quality of work. However, the automation in the construction industry has still depended on the labor input through the use of the construction equipment because of unstructured and dynamically changing work environment and the large handling capacity required by the heavy weight of construction materials. The development of construction robots as well as semi-automated construction equipment has been increased with the rapid development of IT (Information Technology) along with the needs of the construction automation recently. In particular, the development of earthwork automation system is on its way actively because of the highest demands among construction equipments. Now, the development of the IES for the earthwork automation is on progress since 2006 as a part of Construction Technology Innovation Program of Ministry of Land, Transport and Maritime Affairs in Korea. The development of the IES is being performed by the industrial, academic, research institutions of 16 in total such as Hanyang University, Korea Institute of Construction Technology, Korea Electronics Technology Institute, Korea University, Inha University under the principal investigation of DooSan Infracore Co., LTD. IES has a research plan of five years by organizing three research teams after dividing their roles on the basis of three detailed core technologies. The third-year research of the development of the IES is under way now. The system development for remotely-controlled excavator has been almost completed based on the element technical analysis, algorithm and the system structure design which is the first-year research content. The development is kept up on the basis of the close technical cooperation among each division until the fourth year, and the development of the IES will be completed at the end of 2011 by finishing the fifth-year research after going through the performance evaluation of the system.

The final purpose of this project is to improve the productivity, quality and safety in the earthwork, and to develop an intelligent unmanned robotic excavator to be able to overcome the lack of skilled workers and manpower. To do so, TPS, which can establish an efficient working plan by analyzing the geographic information acquired through 3D scanner, should be developed. After compiling a database of a skilled operators' heuristics, TPS is able to devise an efficient work plan.

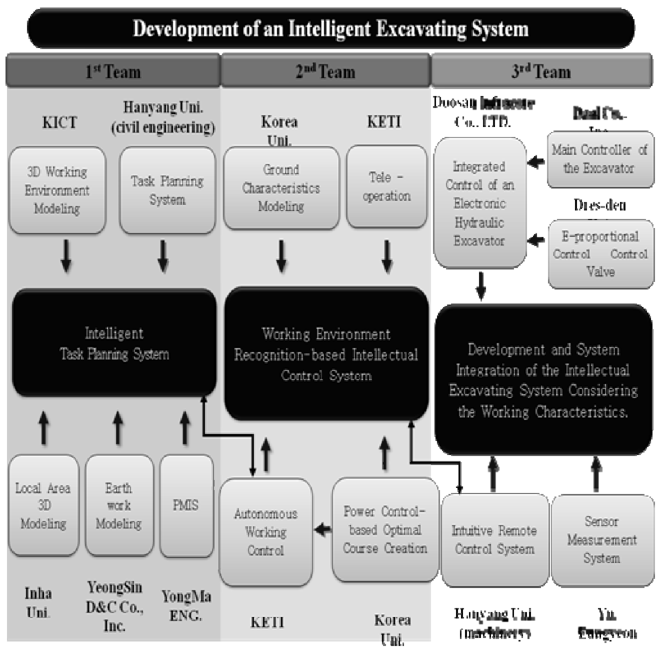


Figure 1. IES research system

1.2 Scope and Methodology

This research collected and compiled heuristics which is skilled workers' experiential knowledge of the earthwork. On the first-year research for the development of TPS, the algorithms for module development was devised based on the heuristics. Task planning simulator to plan and visualize the earthwork process with the virtual reality-based simulation was developed in the second year. The third-year research now is focusing on the integration of TPS with other modules of the IES such as sensors and equipment controller.

The scope of this paper is to present the detailed functions of TPS and to examine the characteristics and the advantages of TPS developed through the skilled operators' heuristics and equipment controller.

2. Intelligent Excavating System (IES)

IES is an excavator-based robot. The excavator robot senses the whole work environment in real-time through sensors and devises the optimal working plan based on the data of the earthwork design and the work environment. On the basis of this, the robot body performs the movement and the work through the intelligent self-control system.

IES is built upon three core technologies. The first technology is 'the intelligent task planning system', which is the development of the work plan creation system to play the core brain role of the IES. The second technology is 'work environment recognition-based intelligent control technology' which is the development of the technology related to the autonomous control as the robotic excavator creates the optimal path for the manipulators. The third technology is 'the development and the system integration of an IES considering the work characteristics', which develops the robot's body and hardware and integrates and manages all the developed system modules. The Fig 2 shows the detailed core technology.

3. Task Planning System (TPS)

TPS is the technology to create the working plans by granting the earthwork and the superintendent's knowledge to enable IES to perform the optimal earthwork plan on the basis of the virtual working environment in a computer identical to the practical working environment updated real-timely on the basis of the sensor data.

TPS performs the tasks related to the whole plan creation on the excavating works such as area division, the optimal platform creation, the sequential creation of works, the optimal excavator movement path plan and creation

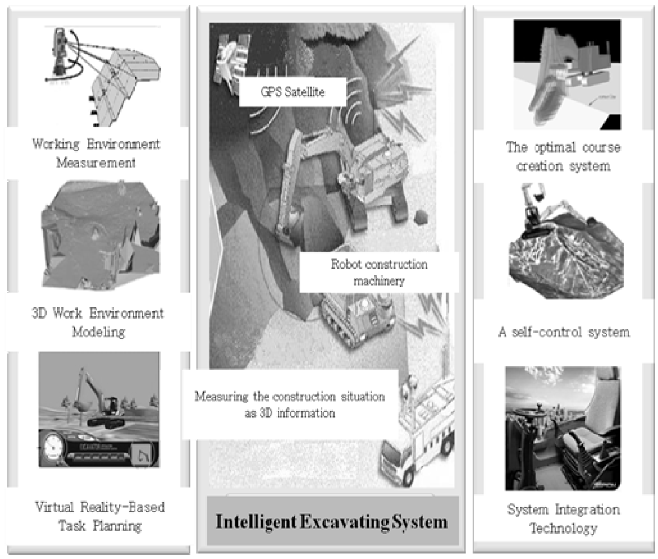


Figure 2. IES's detailed core technology

between the platforms, and the quality control of the working contents, and compiles a database of the work information through the linkage to Project Management Information System (PMIS) which is the construction management module. Also, it is applied as a monitoring system by building the working contents through virtual reality simulation. There are two technologies to acquire the working environment; global sensing and local sensing. Global sensing means the process to map the whole geographic data of the work through laser scanners, and local sensing means the process to map the data for the local area where becomes the object when the excavator is located at the platform to work.

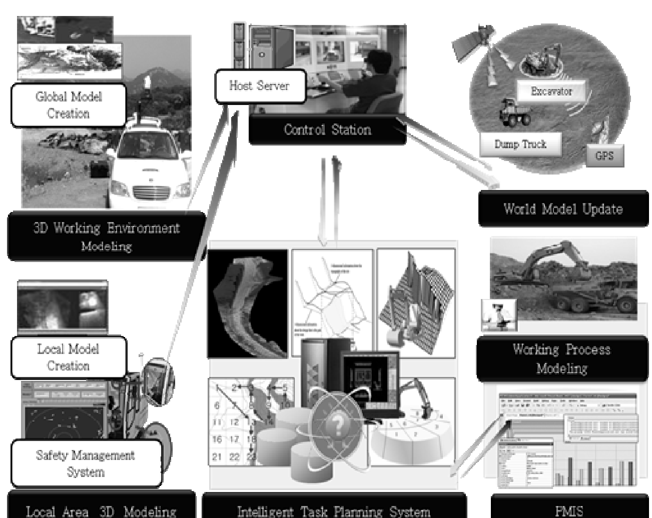


Figure 3. Intelligent TPS Development

4. Development of TPS Module using heuristics

TPS devises the working plan by each module algorithm real-timely by receiving the geographic data through global and local sensing after inputting the work environment data, the dimension information of the robotic excavator in advance. The plan for the work is devised after dividing the target terrain into Global Area, Unit Work Area, Local Area and Local Package. Global Area means the components dividing the whole area for the excavator to work, Unit Work Area means the components fixing the area considering a certain direction and continuity which are the work path characteristics of the excavator, and Local Area means a certain standard of area where the excavator will work after being located at the platform, to be created after Unit Work Area division. Local Package means the components dividing the local area by considering the plan and the characteristics of each work path of the bucket when the excavator works.

The earthwork plan and the earthwork amount of whole work are calculated through the comparison analysis between the data of the real and the designed geographic data. The work order is given after dividing areas at each stage, and the movement path creation of the excavator and the work are performed based on the platform to be located for the excavator.

The Fig. 4 shows the contents on the module and the performance process of TPS designed by applying heuristics.

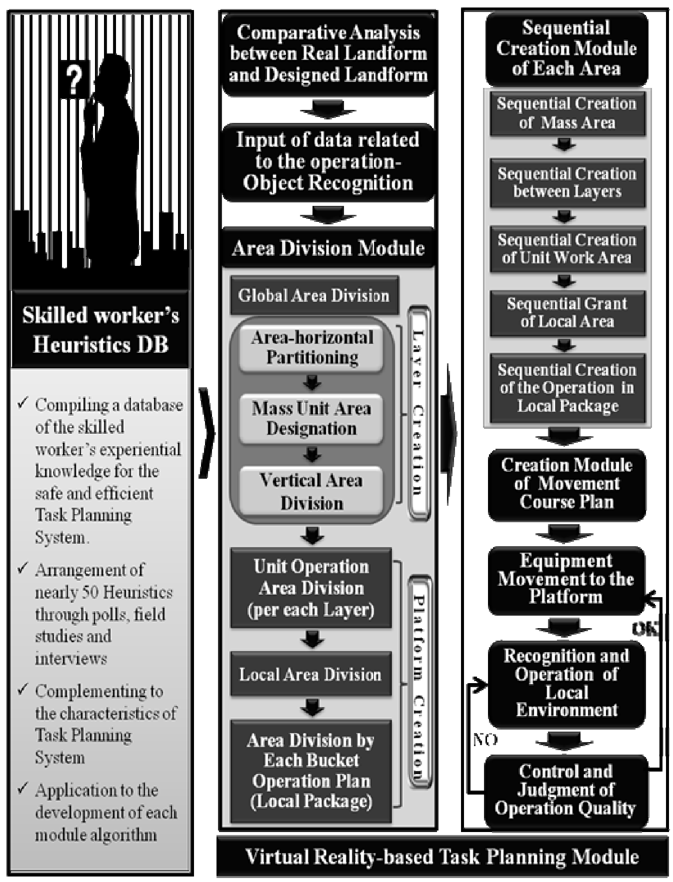


Figure 4. TPS Module & Process

4.1 Division of Global Area

4.1.1 Horizontal Division

This is the process to divide the geographic data horizontally on the basis of the average excavating depth that is efficient most and that considers the safety at the excavating point suitable for the excavator dimension. A skilled operator reduces the risk of a skid by considering the bearing capacity of the ground, and performs the work by making the cutting depth 2~3m deep of the optimal vertical excavating depth according to the excavator dimension at one point, and carries out the work from the high to the low topography in the ground height.

So, as the basic phase or devising the efficient work plan of the excavator, it created layers through the horizontal division of the area according to each phase and district.

4.1.2 Designation of Mass Area

In case of the earthwork field having various peaks, time and cost is saved by minimizing the travel distance of the excavator, and change of direction is saved by applying the work order according to the area after dividing into the mass unit for the whole area.

Global area division performs the work from the upper layer after dividing the geographic data according each layer. Therefore, the location in stages of the excavator on the earthwork field carries out the work by stages from top to down for the efficiency of the excavation and the transport, and the efficient work plan devising can occur if carrying out the work from the upper layer unconditionally without considering the mass unit area.

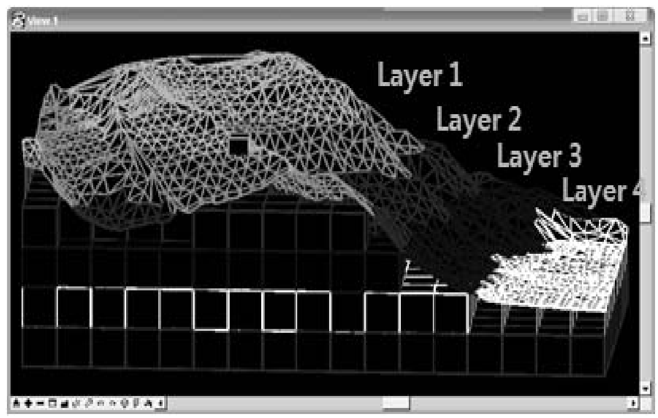


Figure 5. Horizontal Division of the Area

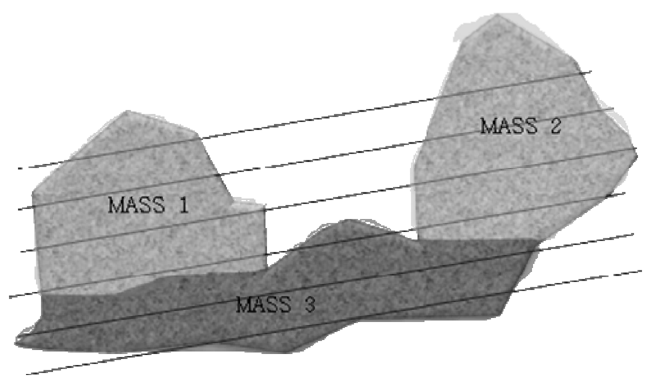


Figure 6. Mass Unit Area Designation

4.1.3 Vertical Division

This is the process to divide the area by the similar works of each geographical feature, and divided according to the earthwork plan. This enables one to consider it when working out the work plan by designating the location of an obstacle or an intrusion by the district to be preserved having been recognized in the geographical information. One secures the transfer road and the working space by removing the intrusion that is not worth its keep and that can be removed by the excavator, and one designates it as a conservation district by recognizing the intrusion to be preserved as the work-excepting section for evading or selecting a detour.

4.2 Designation of Unit Work Area

Unit Work Area is designated by considering the working direction and continuity for the work object layer created through global area division. A skilled operator of the excavator performs continuously with a certain direction until backing is infeasible. This can simplify the work plan and can devise the efficient work plan through the minimization of the movement distance of the excavator and the direction changing. Also, operators perform the work backing the excavator so it becomes well-drained naturally by the gravity to treat the flowing groundwater and the surface water affects on excavating.

The operator decides the spot that has the minimum movement distance and that secures the safety as the next target spot even when moving the distance between the work areas. When designating Unit Work Area, considering obstacles is efficient when beginning the work in the distance place from the place to reduce the damage and to consider the difficult work degree, and enables the platform to be located by considering the corners, obstacles of the layer geographical feature.

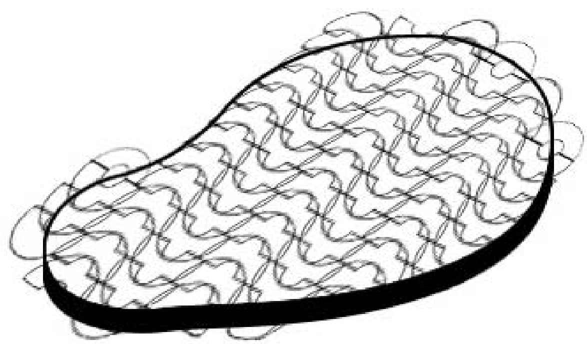


Figure 7. Unit Work Area Designation

4.3 Division of Local Area

This is the process to divide Unit Work Area into Local Area. Platform means the spot where the excavator is located, and the standardized area where the excavator works after being located once as a platform unit area, that is, as Local Area. A certain location of platform position is decided automatically as dividing Unit Work Area into Local Area. The skilled operators perform the excavation as keeping up the angle between boom and arm to be 90~110° for efficient working, and can improve the work efficiency by designating Local Area so that the horizontal rotation angle of the boom is to be applied within 90° when loading the truck. When designing Local Area, the area for gathering the earth and securing the safety is set by having the spatial room of a certain radius. This can reduce the excavators' idle-time and can improve the work efficiency by preventing the inversion due to the loss of the track ground capability and by facilitating loading trucks.

Local Area is platform unit area divided by a certain rule considering the trajectory of bucket and kinematics. The designed geographical feature including each functional and each unit areas organizing Local Area is defined as Local Package. Local Package standards are affected by the specifications and other working environment conditions. Local Package Algorithm was developed by considering these conditions. Local Package Algorithm was made to calculate standards according to Algorithm if inputting variables being required respectively, and U.I (User Interface) for users' convenience was developed.

The inputted variables and designed Local Package through this supply the division standards that become the standards when performing the area division module, and are applied also to the impact data in performing other modules after U.I is configured on the initial screen before the performance of TPS modules.

First of all, the user inputs the track length, the vertical optimal excavation scope, the maximum horizontal excavation scope, the height to the center-joint which are the excavator dimension inputted to the earthwork as the input data of U.I. The optimal horizontal excavation scope length (O.L) considering the excavating depth and the excavator's height by the inputted data, the radius of gathering the earth and the safety guarantee section length (S.L) are to be calculated. The horizontal excavation angle applies 180° by considering a skilled worker's heuristics and the excavating efficiency. The following figure is the plan of the Local Package made by applying AutoCAD.

4.4 Optimal Platform Location Setting

When arranging Local Area and Unit Work Area, the reiteration occurs because the shape of Local Area performs the excavation in the closed circular arc form according to the excavators' characteristics. It is possible to excavate efficiently and economically by removing the unnecessary excavating plan. The number of platforms for the soil- cutting area is decreased by reducing the unnecessary movements and direction changes only when creating the optimal platform location where the reiterations are minimized on the vertical and horizontal way. Also, Local Area standards are changed according to the excavator dimension, the quantified algorithm type is required for an Intelligent TPS to decide automatically the platform location. Algorithm has been developed to calculate the optimal platform location, and verified the minimization of the reiteration by applying AutoCAD.

5. Inspection of TPS with VR Simulation

So far, this paper has explained the modules of TPS, and heuristics applied to organizing the modules. Each module performs its role in series or in parallel and displays its function having an organic relationship mutually, not being

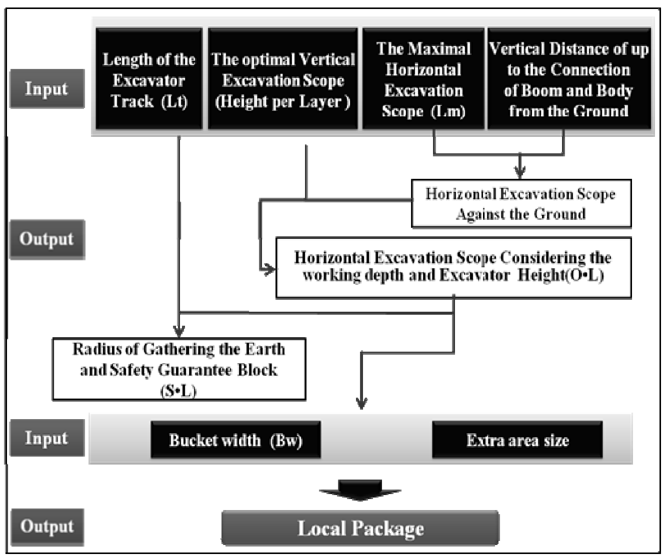


Figure 8 Local Package Algorithm

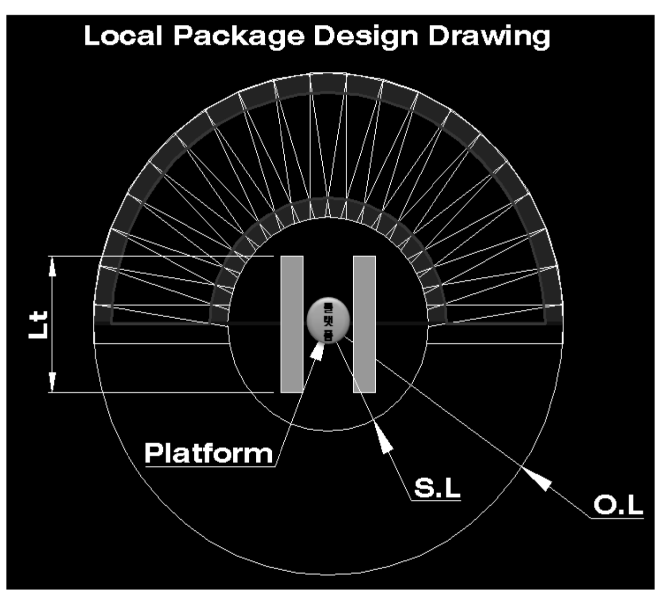


Figure 9. Local Package Design Drawing

independent. Through the area division practically, the sequential creation of works is created naturally. This affects on even the path planning of the excavator bucket and the excavation control. The designated platform supplies the creation and the path plan, and the location for the excavator. Also, the area division process for each bucket supplied the specific unit for the optimal excavation by being designed after reflecting the optimal excavating plan.

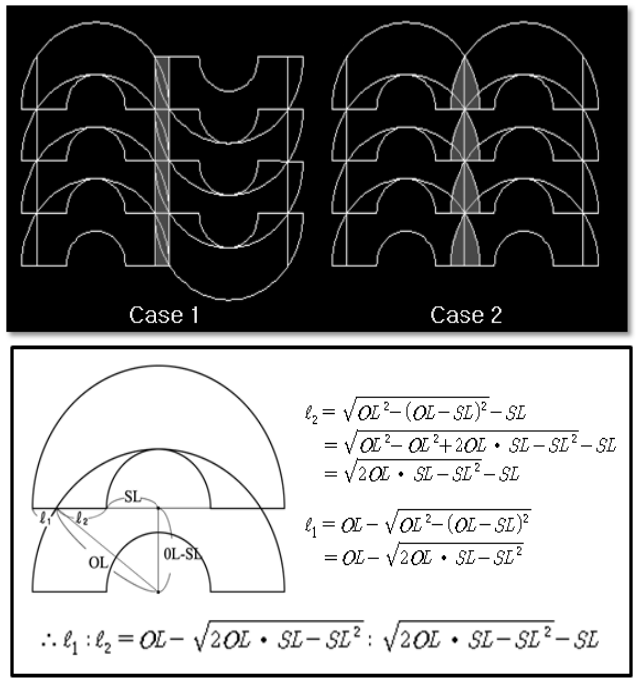


Figure 10. Platform arrangement and reiteration spot calculation considering the minimization of reiteration

The simulation was performed by applying task planning simulator for checking the contents of each module function organizing the intelligent TPS with the system configuration.

First of all, after acquiring the geographic data to be the simulation object, it was changed into solid model without material parameters considering the capacity of the data flow.

Task planning simulator creates Local Package standards by inputting the data necessary in designing Local Package. The depth of the horizontal area division was determined 2.5m, and a layer was created by the horizontal division module. The created layers were changed into solid object models respectively through mesh configuration, and it was designated as Unit work Area creation object after selecting the highest layer. The layer was divided based on the object local model of the calculated 3D solid, and this aims at the changes of the geographic data in three dimensions according to the practical operation of robotic excavator. The location of robotic excavator is displayed on the map after being transmitted through GPS data. The changes of the geographic data through the operation progress are updated on the global geographic data by receiving the formation of the geographical feature changes through the local sensing, and form the condition similar to the figures of the real field by being applied to the virtual simulation of TPS.

In the progress of the simulator, as the operations of the upper layers are progressed, the operation plan was devised simply as the geographical feature became gentle. Also, because the robotic excavator has difficulties in working with a certain direction and continuity if arranging after finding a positionable location by considering the shape of the geographic data when designating the platform location in case of the designation of unit work, the operation plan considering the presence or not of the evenness and the hardening operation possibility of the expected platform spot will be created.

6. Conclusions

Though each country all over the world has performed incessant research and development for the automation of the construction industry, it has run into numerous brick walls such as the continuation of non-repetitive operations, the lasting occurrences of uncertain incidents, the subjugation of the natural environment, the necessary elements, technological shortage, compared to other industries. To solve these

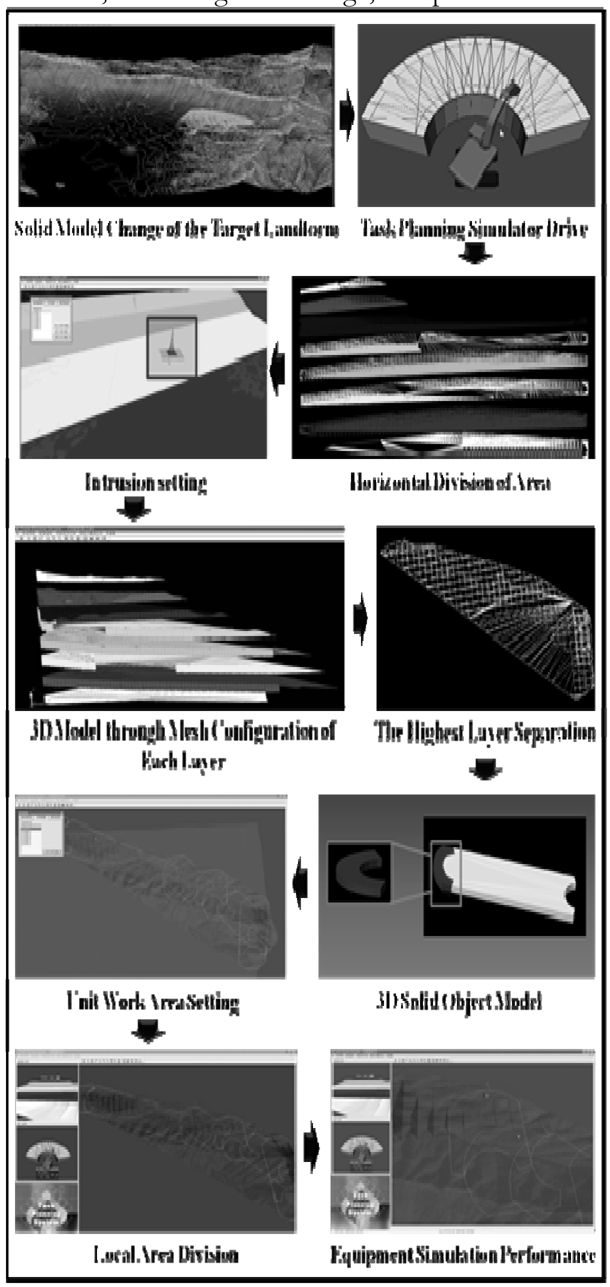


Figure 11. Task Planning Simulator

problems, it should be equipped with the necessary element technologies such as the development of the equipment corresponding to the purpose of the automation, and more exact and rapid sensing, and computer technology. All these things such as the experiential knowledge, the plans to cope with dangerous incidents, the past's successful data to improve the construction productivity for the skilled workers to perform more efficient operation in construction industry now are essential data. In particular, if granting it to the automation system after arranging the skilled workers' heuristics systematically and faithfully in the implementation of the automation through the development of construction machinery, it can bring about numerous profits such as the prevention against various disasters, efficient operations, the improvement in productivity and quality, and will be able to minimize the trials and errors accompanying in the process of achieving the automation. This research improved the efficiency and the safety by applying it to the module configuration through the algorithm development of each stage after arranging the skilled workers' and the superintendent's heuristics being required over the whole of TPS development which is an intelligent plan creation system among the IES development. It is necessary to seek the applications and the plans of various logics to grant artificial intelligence in the future, and the module that trials and errors are fewer and exquisite will be developed only when the development of close algorithm becomes a precondition.

References

[1] Kim, Jeong-Hwan., Lee, Seung-Soo., Park Jin-Woong. and Seo, Jong-Won.(2008). "excavating simulation for earthwork equipment applying Microstation.", 34th KSCE Conference 2008.
[2] Lee, Seung-Soo., Back, Kyung-Geun., and Seo, Jong-Won.(2008). "A study of estimating optimal local area by excavator's standard.", 34th KSCE Conference 2008.
[3] Lee, Seung-Soo., Kim, Jeong-Hwan., Park, Jin-Woong., and Seo, Jong-Won.(2008). "Multi Compatible Excavating System for effective earthwork automation.", 34th KSCE Conference 2008.
[4] Lee, Seung-Soo., Kim, Jung-Hwan., Kang Sang-Hyeok and Seo, Jong-Won. (2008). "Development of Task Planning System for Intelligent Excavating System Applying Heuristics." KSCE Nov. 2008, Vol. 28, Issue 6D.
[5] Lee, Seung-Soo., Kim, Kyoung-Nyun and Seo, Jong-Won.(2008). "Analyzing effect factor in earthwork and skillful excavator operator's behavior.", 34th KSCE Conference 2008.
[6] Lee, Seung-soo, Park, Jin-woong, Seo., Jong-won. and Kim, Sung-keun.(2008). "Decision of Optimal Platform Location Considering Work Efficiency-Optimization by Excavator Specification-", KICEM Conference 2008.
[7] Park, Jin-Woong., Lee, Seung-Soo., Kim, Jeong-Hwan. and Seo, Jong-Won.(2008). "Development of Task Planning System Simulator for Intelligent Excavating System.", 34th KSCE Conference 2008.
[8] Singh, S. and Simmons, R. (1992). "Task planning for robotic excavation.", Intelligent Robots and Systems, Proceedings of the 1992 lEEE/RSJ International Conference on, Vol. 2, No. 7-10, pp.1284-1291.
[9] Seo, Jong-Won., Kim, Young-Wook. and Jang , Dal-Sik(2007). "Development of Intelligent Excavating System -Introduction of research center." KICEM. 2007, pp.921-924.
[10] Seo Jong-Won., Kim Young-Wook., Lee, Seung-soo ., Jang Dal-Sik(2008). "Development of Intelligent Excavating System -Introduction of research progress" KICEM. 2008, pp.184-192.

Development of a Control System for a Multipurpose Road Repairing Machine

Pekka Kilpeläinen[1], Mika Jaakkola[2] and Pauli Alanaatu[3]

[1]VTT Technical Research Centre of Finland, P.O. BOX 1100, Oulu, FIN-90571, Finland; email: Pekka.Kilpelainen@vtt.fi
[2]Destia Ltd, P.O.Box 444,Oulu, FIN-90101,Finland;email: mika.jaakkola@destia.fi
[3]Destia Ltd, Turuntie 207,Espoo, FIN-02740,Finland;email: pauli.alanaatu@destia.fi

Abstract

In this paper an automatic control system for a multipurpose road pavement repairing machine (ROADMOTO) is introduced. ROADMOTO machine is equipped with asphalt milling drum and two asphalt spreaders. The old wearing course of the road is heated and milled. Asphalt spreader in the middle of the machine is used for spreading the crushed old pavement. Asphalt spreader in the back of the machine is used for spreading the new asphalt mass on top of the old layer. Until now most functions of the machine have been manually controlled. The goal is to achieve cost saving and better work quality by using automation. Before the actual repairing work a GPR (Ground penetrating radar), a profilometer or laser scanning techniques are used for collecting information about the road. Designing of repairing tasks is based on the collected data. During a design phase a repair design file is created. ROADMOTO machine is equipped with a GPS positioning unit and the repair design file can be used for automatic control of road repairing operations. The control system also offers a manual control mode as well as automatic height and slope control modes. This ensures flexibility, because the user can choose control mode that best suits for the situation. The control system uses CAN bus as sensor and valve interface and hydraulic actuators are closed loop controlled to achieve high control accuracy. The concept from the data collection and design to the automatic machine control is presented as well as the developed prototype system and results from the first tests.

Keywords: Automation, pavement repairing machines, hydraulic control systems

1. Introduction

It has been estimated that as much as 85% of Europe's road construction projects today include different repairing and rehabilitation operations. Automation is one of the modern means for improving process efficiency and product quality in road construction as well as in road maintenance. The benefits of automation will be produced through the entire construction process. Automated and model based process means exploitation of developed design-, control- and positioning systems in different phases of road construction process.

The process of data flow in road repairing and rehabilitation construction from the automation point of view consists of 1) initial data collection and problem diagnosis, 2) repairing and rehabilitation design, 3) site operations including machine control operations, and 4) quality control actions.

The initial data collection consists of survey methods providing the basic information for rehabilitation design are as follows: a) GPR (ground penetrating radar) for thickness surveys and detecting reasons for damages, b) FWD (falling weight deflectometer) for stiffness measurements of structural layers and subgrade, and c) profilometer or laser scanning techniques to collect data from the road surface. These techniques need accurate positioning systems in order to produce data for precise 3-D road models. The rehabilitation design and machine control models can be processed with special cad tools made for that purpose.

In this paper a control system for a multipurpose road repairing machine (ROADMOTO) is presented (Figure 1). ROADMOTO machine is equipped with asphalt milling drum and two asphalt spreaders. The old wearing course of the road is heated and milled. Asphalt spreader in the middle of the machine is used for spreading the crushed old pavement. Asphalt spreader in the back of the machine is used for spreading the new asphalt layer on top of the old layer. So two layers of asphalt are done at the same time and old crushed asphalt is used on the bottom layer.

The idea of the machine is to use old heated and crushed pavement as much as possible for correcting the worst defects of the road and to minimize the use of new pavement. Typically new mass for correcting the road geometry is needed about 20 kg/m². It is estimated that with automation this can be reduced 50% - 100% and overall cost saving is about 15%. In this case cut and fill operations are done according to a road repairing model.

ROADMOTO machine is also equipped with a mixing drum and it can be used also as an asphalt remixer machine. In this case old crushed pavement is mixed with the new asphalt mass in the mixing drum. The spreader in the middle of the machine is not used in the remixer work.

Figure 1. Multipurpose road repairing machine (ROADMOTO).

2. Surveying and Modeling

Surveying and modelling phases (Figure 2) are important when designing repairing model for road site and applying automation for the repairing of the pavement.

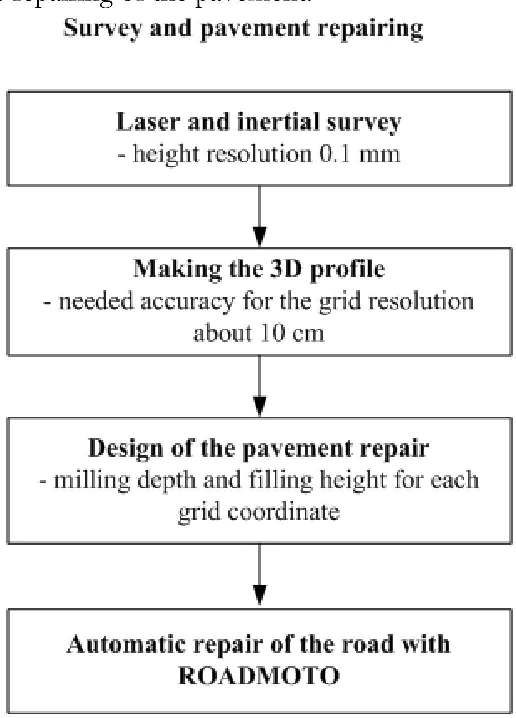


Figure 2. Process to survey and repair the road.

2.1 Survey and 3D model

Surveying of the road is done by special measurement system mounted in a car (Figure 3). Measurement system contains laser, full inertial and GPS devices. System collects all the data from the sensors to the raw data file during the measurement. Next phase is the processing of the collected data. Transversal and longitudinal road profiles are combined in to one 3D surface of the road.

Figure 3. Surveying system.

Accuracy for the 3D model in vertical direction is provided by high resolution laser units (0.1 mm). In horizontal direction grid resolution is possible to choose to be high enough for terrain modelling applications. Ground penetrating radar can be used, if information about road structure e.g. pavement thickness is needed.

2.2 Design for repair

Design for repair is done with terrain modelling application. Design is done with the accuracy level specified. In the designing phase of the repair process it has to be known what kind of repairing work will be done and what the target of repair is. Different kinds of things of the pavement condition has to be surveyed and decided how to fix them. There can be e.g. edge drops, potholes or deep rutting. Various kinds of road damages need special actions when repaired.

The amount of asphalt used is optimized in the designing phase. E.g. there are sections where it is possible to mill more than needed for repairing and there will be sections where the crushed asphalt will be used for filling. GPR data is useful in optimization.

When designing is ready the plan is inserted in to the automated pavement repairing machine ROADMOTO. Repairing file is created and ROADMOTO's control software can read it as input data. Input data file defines the milling depth and filling height for each grid coordinate. The input data file covers whole construction site. If spreader is used in the repairing work, the spreader height level for each grid coordinate is given and it is related to milled surface.

3. Machine Control System

3.1 Main functions

The development work of the control system started with analysing different working situations during operation of the machine. Two main situations are:

-No surveying and modelling is done beforehand. This is typical situation in small roads and work sites.

-Surveying and modelling is done beforehand and a repairing model is available for automatic machine control. Surveying and modelling are typically done only in main roads and large work sites.

The machine control system should be applicable for both of these situations. Because the circumstances vary on a work site considerably, choosing between manual and automatic operation should be as flexible as possible.

The automatic control system is used for controlling the milling drum and the asphalt spreader in the middle of the machine. These are referenced below as the cutter and the spreader. Asphalt spreader in the back of the machine is controlled manually and is not connected to the automation system.

Both the cutter and the spreader are moved by two hydraulic cylinders (Figure 4). The variables that the automation system controls are cutting depth (or height) h_1 and h_2 and slope θ of the cutter (or spreader).

Ultrasonic sensors are used for measuring the actual height h_{1m} and h_{2m} relative to the road surface and inclination sensor is used for measuring actual slope a_m.

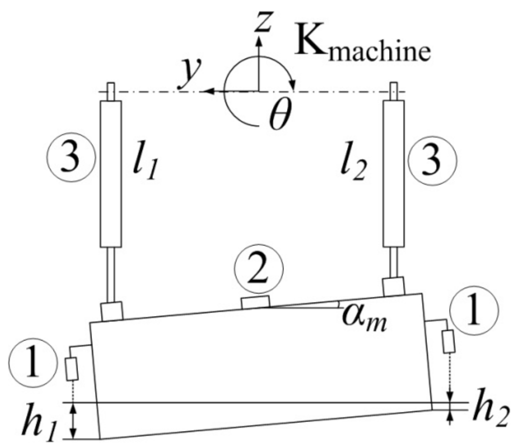


Figure 4. Hydraulic cylinders of the cutter and the spreader

The ROADMOTO machine weights 40 tons, length is 16 m and wheelbase is about 8.5 m. The machine is equipped with rubber tyres in the front and air tyres on the back. Compression of the tyres and the torsion of the frame also affect the position and orientation of the cutter and the spreader.

Operating modes

To allow most flexibility to the user, the control system was designed so that user can set a control mode individually for each of the four cylinders. Different control modes are marked by letters J, U, K and M. Control modes are:

- *J (Joystick mode):* Full manual control using a joystick.
- *U (Ultrasonic mode):* Cutting depth and height control using ultrasonic sensors (Figure 4 number 1).
- *K (Slope control):* Automatic slope control using an inclination sensor (Figure 4 number 2).
- *M (Model control):* Automatic slope control according to the repairing model.

Changing between control modes can be done on-the-fly during operation. Because the user can choose from four different modes for both cylinders, there are 16 possible combinations of control modes. Although the number of combinations is quite high, the control logic is easy to learn and offers flexibility.

3.2 Control method

Controlling the motions of the cutter and the spreader, a Cartesian control method is used (Figure 5). The benefit of the Cartesian control is that different Cartesian values can be controlled independently (height of left end h_1 and height of the right end h_2 and absolute slope a) according to the selected control mode (m_1 and m_2). Joysticks (j_1 and j_2) can override automatic control.

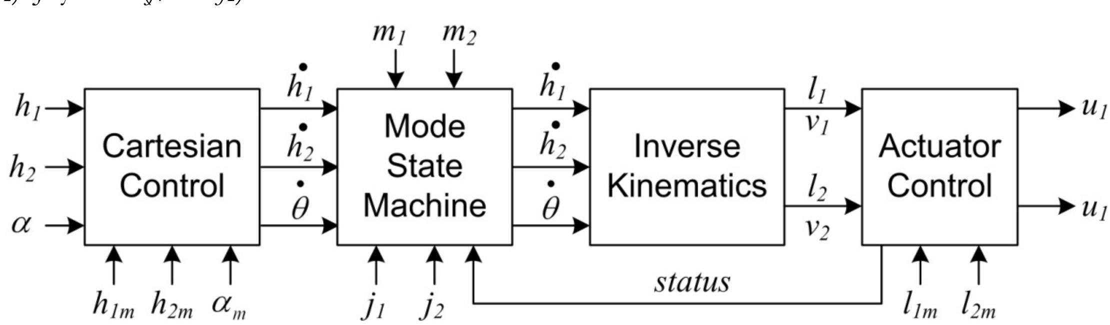


Figure 5. Cartesian control method.

Outputs from the Cartesian control are Cartesian velocities $\dot{h}_1$, $\dot{h}_2$ and $\dot{\theta}$. From these set values for the positions (l_1 and l_2) and velocities (v_1 and v_2) of the cylinders are calculated. Positions of the cylinders are then closed-loop-controlled. *Status* of the actuator control is used as feedback. If delay of cylinders exceeds limits, Cartesian velocities are limited.

In this case the mechanism that is controlled is quite simple. Cartesian control based on solving inverse kinematics gives more advances when controlling more complex mechanism and when manual control is demanding and high accuracy is needed. Some applications are co-ordinated manual control and automatic control based on real time positioning and 3D models of the surface, e.g. road. Some examples are road grader [Kilpeläinen 1999], stabilization cutter [Kilpeläinen 2004] and excavator [Makkonen 2006].

3.3 Hardware and software

The control system architecture is a combination of centralized control and distributed IO. The control system consists of two user interface modules (one in both sides of the machine, Figure 6) and four IO modules, which are connected together using a CAN bus (Controller Area Network). Analogue sensors, such as ultrasonic sensors, are connected to the CAN bus via IO modules. Devices supporting CANOpen protocol (e.g. directional valves and inclination sensors) are connected directly to the CAN bus. Use of CAN bus reduces the amount of wiring considerably.

Control software is implemented in a centralized way, which means that all the control tasks are handled by the user interface module. Both of the user interface modules can work as a CAN master and handle control tasks. The benefit of this is the duplication of control devices, which increases reliability. The user interface module is based on a 16 bit digital signal processor from Freescale. That offers 60 MIPS performance for control tasks.

Figure 6. User interface module.

User interface module consists of joystick for manual control and switches for changing operational modes and set values (e.g. height can be set in 1 mm steps). Current settings are presented in a graphical user interface.

3.4 Model based control mode

In the model based control mode the ROADMOTO machine is equipped with PC computer and RTK (Real Time Kinematic) GPS receiver. PC computer handles positioning and reads set values from the repairing model according to the position of machine. These values are transferred to the machine control system via CAN bus. If the user has selected model control mode (M mode), height of cutter (or spreader) are controlled automatically according to the values from repairing model.

4. Prototype Test

The system was taken in use in June 2008. During summer 2008 the system has been in use for about two months. In October 2008 a test was arranged in order to test the whole concept from the surveying and modelling to the automatic machine control.

4.1 Test site

Test site is typical road in south Finland (Figure 7). Annual average daily traffic in the test site is approximately 2000 vehicles. There were no problems during the survey or making the 3D profile due to low traffic.

Condition of the test site was good and the main reason for repairing was the rutting. Some sections had edge drops, but only 20 – 30 cm from the edge was damaged.

Repairing work with ROADMOTO was done during daytime. All the ROADMOTO sensors were mounted and raw data was collected during the test. Both full 3D profiling and automatically controlled repair work were successful. The test site considered total 26 000 m^2 of repairing work.

Figure 7. Test site.

4.2 Testing of the control system

During the tests different operating modes of the system were tested. Typical situation was that ROADMOTO machine was driven along the left lane of the road. In this case ultrasonic control was used on the right side of the machine (centre of the road) and manual control was used on the left side (road border). Ultrasonic control is not applicable if the surface of the road is very uneven, which is typical for road border.

In figure 8 the functioning of the control system is shown when model control (M mode) is used on the right side of the machine. During the 20 second time period in the figure the machine moves forward about 2 m. In the upper graph set value for height h_2 and measured value h_{2m} are shown. In lower graph set value l_2 and measured value l_{2m} of the position of the cylinder are shown.

The closed-loop-control is done over CAN bus. Simple P-controller and feed forward of the velocity set value is used [$u = K_p \cdot e + K_v \cdot v$]. Sampling time T_s was 20 ms. Accuracy of the position control of the cylinders is about ± 1 mm.

Relative accuracy of the height control is about ±5 mm. The drawback of the test was that absolute accuracy of the height control could not be verified properly. Height control is based on ultrasonic sensors. Temperature changes affect to the output of the ultrasonic sensors, because the velocity of sound changes 0.18% / ºC. To compensate this integrated temperature compensation is used in ultrasonic sensors. In the case of heated asphalt air near the surface can be very hot (about 150 ºC). Ultrasonic sensors still work quite well, but the big temperature difference in the air increases noise of the measurement. Also the zero position of the height, that is set during calibration of the system, changes. To overcome this problem development work is still needed.

4.3 Quality measurements

After the repairing work quality measurement was done. One way to measure quality of the road is Internatiol Roughness Index (IRI) [Sayers 1986].

IRI was measured before and after the repairing job (Figure 9). In the Figure 9 is one section where ROADMOTO was doing automated repairing. IRI level 0.8 is good and it is not easy to go under that value just doing repairing of the pavement. High original value between the distances from 600 to 700 was smoothed in the repairing work to be 0.8 mm/m.

5. Conclusions

The ROADMOTO machine equipped with the control system presented in this paper has been in use since June 2008. The repairing process of the road pavement consists of surveying, design of repairing, site operations including machine control and quality control tasks. The developed control system can exploit the design data, a road repairing model, in automatic machine control. Prototype test were arranged to test this concept. All the phases from the survey using a car mounted surveying system to the repair design and automatic machine control were completed. Finally achieved quality was measured.

Although there exists some problems e.g. in the calibration of the machine control system, the presented automatic control method is applicable. It is also important to that the user can choose between manual and automatic features according to the situation. This was also taken into account during the design of the control system and the presented system is also useful in work sites, where no surveying and design is made beforehand.

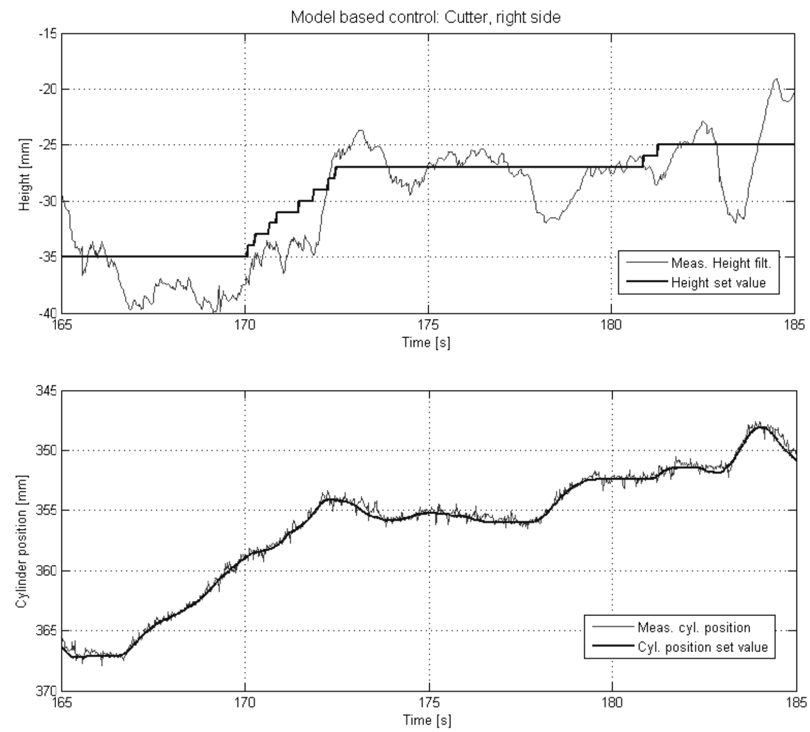


Figure 8. Automatic control of the right side of the cutter.

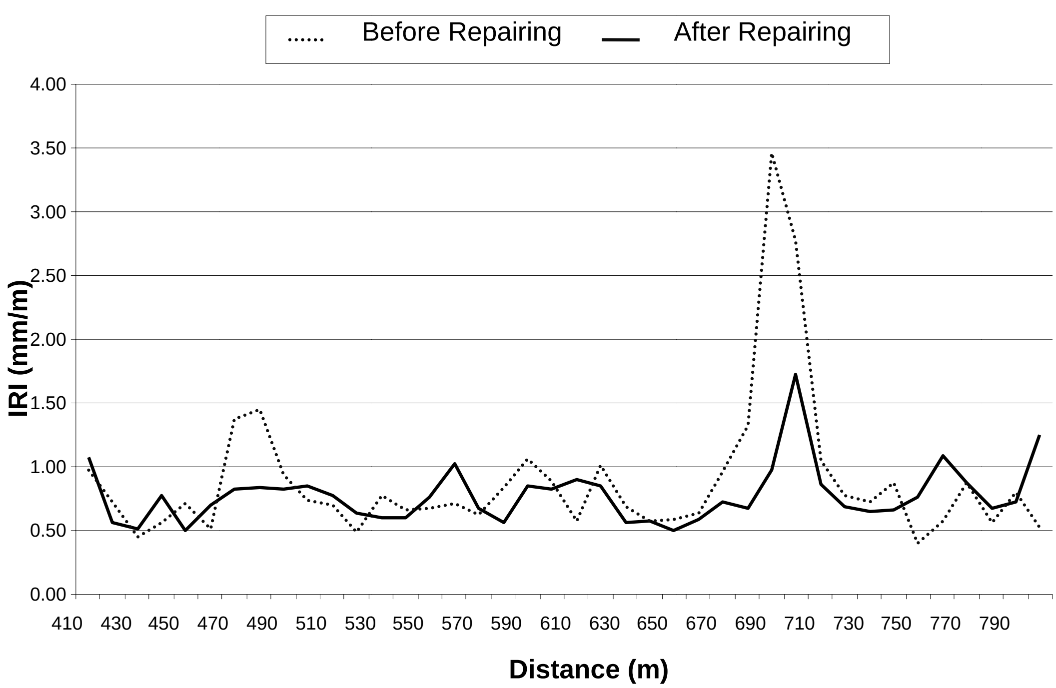


Figure 9. Road quality before and after the repairing work.

During the short period of testing it was not possible to fully evaluate the benefits of the system. The next goal is to take the presented process into common use.

References

[1] Kilpeläinen, P., Nevala, K. (1999). "Design and simulation of a road scraper control using the TeleGripTM simulation program", Proceedings of Scandinavian Symposium on Robotics 1999. Oulu, 14 - 15 Oct. 1999. pp. 232 - 239.
[2] Kilpeläinen, P. Nevala, K. Heikkilä, R. (2004). ”Development of a control system for road construction automation applications”, Proceedings of 21st International Symposium on Automation and Robotics in Construction. Jeju, 21 - 25 Sep. 2004. IAARC, pp. 65 – 70.
[3] Makkonen, T., Nevala, K., Heikkilä, R. (2006). ”3D model based control of an excavator”. Automation in Construction, vol. 15, 5, pp. 571 – 577.
[4] Sayers, M., Gillespie, T., Paterson, W. (1986). "Guidelines for Conducting and Calibrating Road Roughness measurements". World Bank Technical Paper, vol. 46, 5, ss. 1 - 12

Measurement of the International Roughness Index (IRI) Using an Autonomous Robot (P3-AT)

Jia-Ruey Chang[1], Yung-Shuen Su[2], Tsun-Cheng Huang[3], Shih-Chung Kang[4] and Shang-Hsien Hsieh[5]

Department of Civil Engineering & Department of Leisure Management, MingHsin University of Science & Technology, No. 1, Hsin-Hsing Road, Hsin-Chu 304, Taiwan; PH (886) 3559-3142#3873; FAX (886) 3557-4451; e-mail: jrchang@must.edu.tw[1], chenww5828@yahoo.com.tw[3]
Department of Civil Engineering, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei 106, Taiwan; PH (886) 2-3366-4346; FAX (886) 2-2368-8213; e-mail: operaticlife@gmail.com[2], sckang@caece.net[4], shhsieh@caece.net[5]

Abstract

In this paper, we test whether an autonomous robot can be used to measure the International Roughness Index (IRI), a description of pavement ride quality in terms of its longitudinal profile. A ready-made robot, the Pioneer P3-AT, was equipped with odometers, a laptop computer, CCD laser, and a SICK laser ranger finder to autonomously perform the collection of longitudinal profiles. ProVAL (Profile Viewing and AnaLysis) software was used to compute the IRI. The preliminary test was conducted indoors on an extremely smooth and uniform 50 m length of pavement. The average IRI (1.09 m/km) found using the P3-AT is robustly comparable to that of the commercial ARRB walking profilometer. This work is an initial step toward autonomous robotic pavement inspections. We also discuss the future integration of inertial navigation systems and global positioning systems (INS and GPS) in conjunction with the P3-AT for practical pavement inspections.

Keywords: Pavement, Smoothness, International Roughness Index (IRI), Autonomous Robot, ProVAL

1. Introduction

Road roughness, or smoothness, inspections are performed to monitor the pavement conditions in order to evaluate the ride quality of new and rehabilitated pavements. Roughness is closely related to vehicle operating costs, vehicle dynamics, and drainage. The American Society for Testing and Materials (ASTM) E 867 defines roughness as the deviations of a pavement surface from a true planer surface with characteristic dimensions. A pavement profile represents the vertical elevations of the pavement surface as a function of longitudinal distance along a prescribed path of travel (Wang, 2006). Both manual and automatic multi-function profiling systems are continuously being developed and marketed for improved performance.

In this study, we pioneer the use of an autonomous robot (the P3-AT) to perform roughness inspections for project-level pavement management purposes. The P3-AT is able to autonomously collect longitudinal profiles at prescribed sampling intervals and compute the International Roughness Index (IRI). It is anticipated that the robot will replace manually operated equipment for construction QC/QA purposes in the near future. A preliminary test on an extremely smooth and uniform 50 m pavement section shows that the average IRI (1.09 m/km) obtained by the P3-AT is comparable to commercial Australian Road Research Board (ARRB) walking profilometer. We also discuss the future integration of inertial navigation systems and global positioning systems (INS and GPS) in conjunction with the P3-AT for practical pavement inspections.

2. Literature Reviews

Pavement profiling systems started with straightedge devices in the early 1900s. Other simple profiling devices, profilographs, and response type road roughness measuring systems (RTRRMS) were developed in the late 1950s and 1960s. Between the late 1960s and 1980s, highway agencies primarily adopted the profilograph for measuring and controlling initial roughness of new construction pavement. The use of inertial profilometers in monitoring pavement condition increased in the 1980s and early 1990s (Wang, 2006). The aforementioned equipment can be divided into five categories (Perera and Kohn, 2002):

- Manual devices: rod and level surveys, straightedge, rolling straightedge (high-low detector), Dipstick, ARRB walking profilometer, etc.
- Profilographs: Rainhart profilograph, California profilograph, etc.
- RTRRMS: Bureau of Public Roads (BPR) roughometer, Mays Ride Meter (MRM), Portland Cement Association (PCA) ridemeter, etc.
- High-speed inertial profilometers: Automatic Road ANalyzer (ARAN) by Roadware Group Inc., Model T6600 Inertial Profilometer by K. J. Law Engineers Inc., etc.
- Lightweight profilometers: Model 6200 lightweight inertial surface analyzer (LISA) by Ames Engineering, Inc., CS8700 lightweight profiler by Surface Systems & Instruments, Dynatest/KJL 6400 lightweight profilometer by Dynatest Consulting, Inc., etc.

The high-speed inertial profilometer can be fitted to full-sized vehicles to measure the pavement profiles at traffic speed. Most lightweight profilometers currently commercially available integrate using the same hardware used in high-speed inertial models mounted on golf carts or small vehicles. ASTM E 950 defines inertial profilometers as Class 1 to Class 4 according to their sampling interval, vertical measurement resolution, precision, and bias. The commonly used modern devices for profiling such as rod and level, Dipstick, ARRB walking profilometer and most inertial profilometers and lightweight profilometers are all Class 1 devices (FHWA-LTPP Technical Support Services Contractor, 2004).

The high-speed inertial profilometer is commonly used to perform roughness inspections for network-level pavement management purposes, but other approaches (such as the California profilograph, ARRB walking profilometer, and lightweight profilometers) have been specifically developed for project-level pavement management purposes. Furthermore, although high-speed inertial profilometers dominate today's market, their application to construction acceptance testing for new or rehabilitated pavements remain limited due to their high cost and scheduling limitations – the short-length pavement overlay and the tests on rigid pavements, for instance, cannot be performed until after a few days of curing (Kelly et al., 2002). As such, most highway agencies use primarily manual methods for QC/QA purposes of new pavement construction and small-scale rehabilitation projects, with the high-speed inertial profilometer used for extended measurements over time (Baus and Hong, 1999).

Roughness indices are derived from profile data and correlated with road users' perceptions of ride quality to indicate the level of pavement roughness. These include the Profile Index (PI), International Roughness Index (IRI), Ride Number (RN), Michigan Ride Quality Index (RQI) and Truck Ride Index (TRI) (Sayers and Karamihas, 1996). Among them, IRI is the index most widely used for representing pavement roughness. ASTM E 1926 defines the standard procedure for computing the IRI from longitudinal profile measurements based upon a mathematical model referred to as a quarter-car model. The quarter-car is moved along the longitudinal profile at a simulation speed of 80 km/h and the suspension deflection calculated using the measured profile displacement and standard car structure parameters. The simulated suspension motion is accumulated and then divided by distance travelled to give an index with unit of slope (m/km), the IRI. Most highway agencies are using the IRI to evaluate new and rehabilitated pavement condition, and for construction QC/QA purposes (Wang, 2006). The IRI can be reported two ways:

- Single path IRI: Based on a quarter-car model run over a single profile.
- Traffic lane IRI: A composite result representing the roughness of a traffic lane. It is determined by averaging two individual, single path IRIs obtained separately in each wheel-path (at 0.75 m either side of the lane mid-track).

Some surveys of state highway agencies in the United States indicate that about 10 percent (4 of 34 respondents) use IRI to control initial roughness (Baus and Hong, 1999), while about 84 percent (31 of 37 respondents) use IRI to monitor pavement roughness over time (Ksaibati et al., 1999), making IRI the statistic of choice for roughness specifications. The proposed 2002 Design Guide under development by the National Cooperative Highway Research Program (NCHRP) also included IRI prediction models that are a function of initial IRI (IRI0) (Kelly et al., 2002; Baus and Hong, 1999).

The lightweight profilometer has been shown to obtain timely and accurate measurements of pavement profiles and to be significantly faster than profilographs (24 km/hr versus 5 km/hour). However, like the high-speed inertial profilometers, they require operators to perform repetitive, tedious, and time-consuming procedures, their awareness and knowledge of the profiling systems and influencing factors largely determining the efficiency of the measurement. The profilometer, for example, must be operated by experienced inspector at relatively constant speeds and the wheel-path must be consistent between measurements. (Mondal et al., 2000).

This study proposes using robots to complete project-level pavement roughness inspections, by autonomously collecting longitudinal profiles and computing the IRI, thus increasing mobility and accuracy, and minimizing time and labor commitments. With self-controlled and automatic motion, robots can reduce the variation and uncertainty of profile measurements and improve inspection reliability.

3. Autonomous Robot Preparation

In this section, we briefly introduce the preparation of the autonomous robot. An integrated set of vertical displacement sensors (CCD laser), odometers, SICK laser ranger finder, and control laptop are mounted on the P3-AT, which is manufactured by MobileRobots Inc (2008). The P3-AT, which can move up to 3 km/h, is capable of measuring longitudinal profiles using a CCD laser at 15 cm or smaller sampling intervals, from which the IRI can be simultaneously computed using laptop-based ProVAL software.

3.1 Autonomous Robot: Pioneer 3-AT (P3-AT)

Being powerful, easy-to-use, reliable and flexible, the P3-AT used in this study is a highly versatile all-terrain robotic platform particularly suited to pavement inspections. Figure 1 shows the P3-AT (MobileRobots Inc., 2008), equipped with a control laptop, onboard Pan-Tilt-Zoom (PTZ) camera system, Ethernet-based communications, a SICK laser and eight forward and eight rear sonars which sense obstacles from 15 cm to 7 m. The P3-AT's powerful motors and four robust wheels can reach speeds of 0.8 m/sec and carry a payload of up to 30 kg. It can climb steep 45% grades and sills of 9 cm and uses 100-tick encoders with inertial correction recommended for dead reckoning to compensate for skid steering. Its superior sensory system employs laser-based navigation options, integrated inertial correction to compensate for slippage, bumpers, a gripper, vision, stereo rangefinders, a compass and a rapidly growing suite of other options. The bare P3-AT base includes the Advanced Robotics Interface Application (ARIA) software enabling the user to (MobileRobots Inc., 2008):

- make the P3-AT move randomly;
- drive using key or joystick control;
- plan inspection paths with gradient navigation;
- display a pavement spatial map using sonar readings, laser readings or a combination of the two;
- localize using sonar or laser upgrade;
- communicate sensor and control information relating sonar, motor encoder, motor controls, user I/O, and battery charge data;
- test pavement inspection activities quickly with ARIA API (in C^{++});
- simulate pavement inspection behaviors offline with the simulator that accompanies each development environment.

3.2 Laser for Vertical Displacement Measurements: LK-G155

In order to conduct inspections on extremely smooth pavement (e.g. a new construction pavement), the resolution of the sensor signals must be very high. Laser sensors are best suited to this purpose. In this study, a LK-G155 (a CCD laser displacement sensor), as shown in Figure 1, is used to measure the vertical displacement from laser to pavement surface, at 15 cm sampling intervals while the P3-AT is in motion. It is mounted in front of the P3-AT and is situated 15 cm above the pavement. The sensor head specifications of LK-G155 are as follows (Keyence Corporation, 2008):

- Mounting mode: Specular reflection
- Reference distance: 147.5 mm
- Measuring range: ± 39 mm
- Spot diameter (at reference distance): Approximately 120 x 1700 μm
- Resolution: 0.5 μm
- Linearity: ±0.05% of F.S. (F.S. = ± 40 mm)
- Sampling frequency: 20/50/100/200/500/1000 μs (selectable from 6 levels)
- Weight (including the cable): Approximately 290 g

- Light source: Red semiconductor laser with 650 nm wavelength (visible light), Class II (FDA), 0.95 mW maximum output
- Temperature characteristics: 0.01% of F.S./°C (F.S. = ± 40 mm)
- Resistance to vibrations: 10 to 55 Hz, multiple amplitude 1.5 mm; two hours in each of X, Y, and Z plane
- LED display: Green in the centre, within the measurement area is orange lights, outside the measurement area is flashing orange.

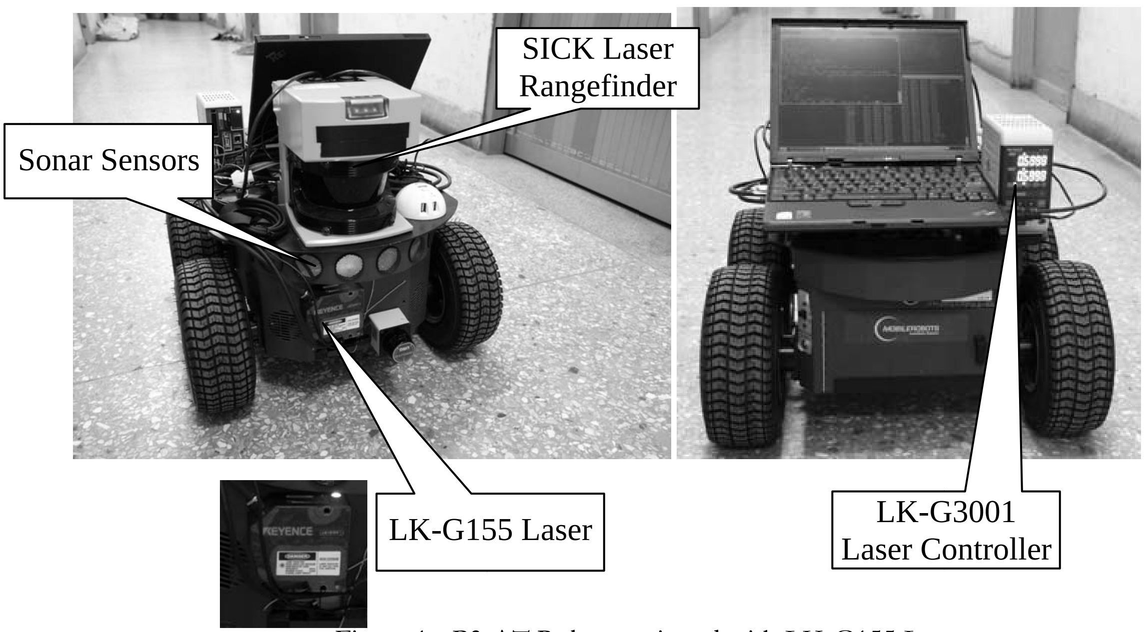


Figure 1 - P3-AT Robot equipped with LK-G155 Laser

Spot size is an important consideration. A small laser spot increases the detector's sensitivity to coarse pavement macro-textures, thus resulting in a profile including high frequency content not relevant to roughness. The LK-G155 uses the wide spot optical system that has high measurement stability. In order to avoid data fluctuations, diffused reflections caused by pavement surface irregularities are averaged with multiple samples to reduce the amount of texture noise at each point of the measured profile. Moreover, this issue does not present a problem because of the effective anti-aliasing process applied to the laser signals.

The digital vertical displacements measured by LK-G155 are displayed and controlled by LK-G3001 controller (as shown in Figure 1). The LK-G3001 controller can transmit the displacement measurements to the laptop via USB at a high-speed 50 kHz. The profile measurements can then be further imported into the ProVAL (Profile Viewing and AnaLysis) software to compute the IRI.

3.3 ProVAL (Profile Viewing and AnaLysis)

The ProVAL engineering software package, developed by the Transtec Group (2008), allows analysts to view and analyze pavement profiles in many different ways. It is easy to use and can perform various profile analyses, including profile editing, standard ride statistics (IRI, RN, etc.), profilograph simulation, rolling straightedge simulation and ASTM E 950 precision and bias. ProVAL is a product sponsored by the US Department of Transportation, Federal Highway Administration (FHWA) and the Long Term Pavement Performance Program (LTPP).

Version 2.73 was installed on the laptop. The LK-G3001 controller imports profile measurements obtained from the LK-G155 laser into ProVAL, which then analyzes IRI for each longitudinal profile as soon as one profile inspection has been finished. Analysts are then able to print a report of the original profiles and of any analyses performed.

4. Preliminary Roughness Inspection Test

In this section, we describe the preliminary indoor test and the comparison of measured IRIs between the P3-AT and an ARRB walking profilometer.

4.1 Test Section and Test Plans

A 50 m straight test path was designated on a smooth, uniform indoor test section, as shown in Figure 2. The test procedure was as follows:

- LK-G155 laser test: The precision and bias of the laser sensor was tested under a static situation before the preliminary test.
- Longitudinal distance test: This test was used to evaluate the precision and bias of the odometer and SICK laser ranger finder. Coupled with the odometer, the SICK laser ranger finder was used to ensure the P3-AT moves accurately along the test path and at a consistent and precise speed.
- Profiling of test path: The test path was used to compare profiles and IRIs between the P3-AT and commercial equipment. Repeated measurements of the IRI were found to evaluate the repeatability of the P3-AT. The ARRB walking profilometer was also used along the same path. The results from each method were then compared.

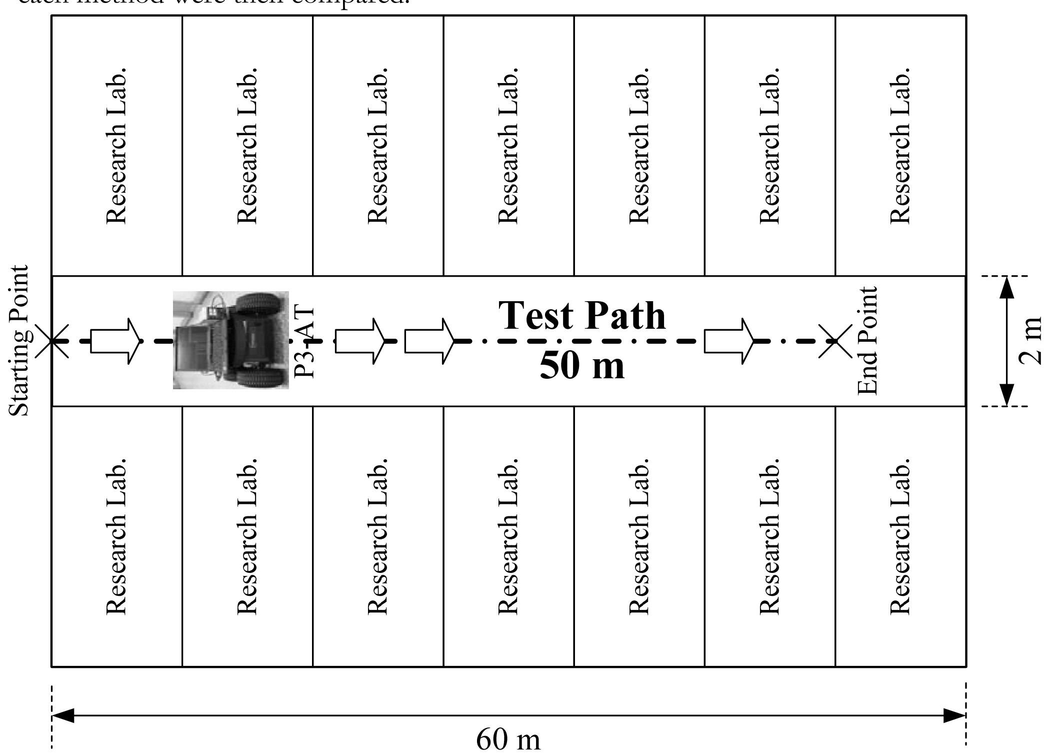


Figure 2 - Test section

4.2 Test Procedures and Test Results

The P3-AT, moving along the test path at a speed of up to 3 km/h, autonomously stops at each 15 cm sampling interval to measure the vertical displacement with the LK-G155 laser. In general practice, sampling intervals range from less than 25 mm to 380 mm. The P3-AT's speed has no effect on the result because the vertical displacement measurement is found when the unit is at rest. The odometer is used to determine the longitudinal distance. The SICK laser range finder is used to ensure the P3-AT follows the test path. No accelerometer is required on the P3-AT, usually needed to compensate for the vertical acceleration of the unit itself, because the test surface is uniformly smooth and level. The commercial device, the ARRB walking profilometer, was then used to obtain another computation of the IRI on the test path. The average IRI (1.09 m/km) from several runs on the test path shows that results from the P3-AT are comparable to the IRI (1.11 m/km) obtained from the ARRB. Figure 3 shows the vertical displacement measurement dataset, in *.ERD format, imported from LK-G3001 controller. Figure 4 displays the original profile of the imported file and the metric system of units has been selected, with distance in meter and elevation in millimeters. Figure 5 shows the effect of IRI filter (with 250 mm filter) on the original profile. Figure 6 shows the result of IRI computation (1.09 m/km).

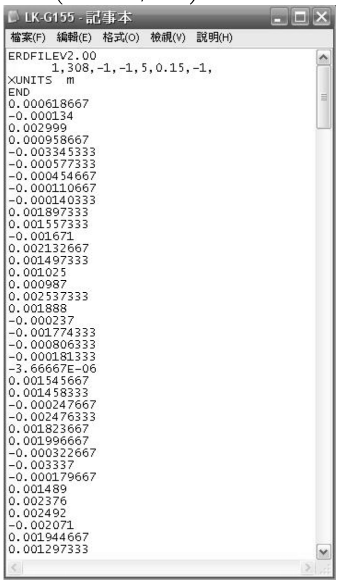


Figure 3 - The vertical displacement measurement dataset in *.ERD format

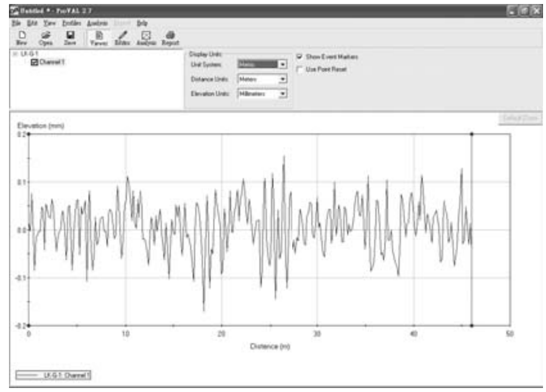


Figure 4 - The original profile

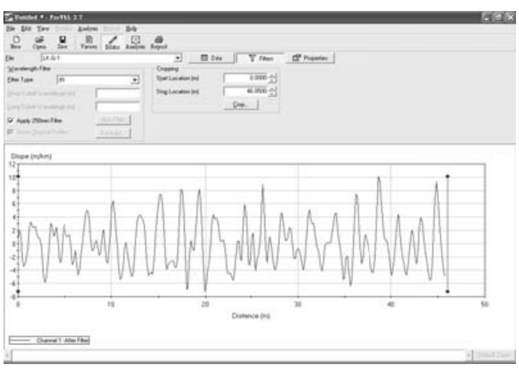


Figure 5 - The filtered profile by IRI filter (250 mm filter)

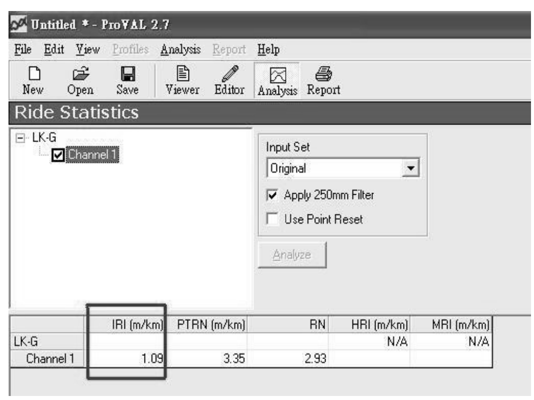


Figure 6 - The IRI computation (m/km)

5. Conclusions and Remarks

In the study, we propose using a P3-AT robot to perform roughness inspections for project-level pavement management purposes. The P3-AT can autonomously conduct the collection of longitudinal profiles using a 15 cm sampling interval and compute the International Roughness Index (IRI). It is anticipated that the use of an autonomous robot will replace manually operated or driven equipment for construction QC/QA purposes in the near future. The preliminary test on an extremely smooth 50 m test path shows that the average IRI (1.09 m/km) obtained by the P3-AT is comparable to a commercial ARRB walking profilometer (1.11 m/km). The inspection architecture of P3-AT has proven to be very reliable. In the future, the test path can be profiled with the P3-AT at different speeds to evaluate its effect on profiles and IRIs.

For the current inspection architecture, there is no accelerometer mounted on P3-AT since the test path is very level, the pavement surface is extremely smooth, and the vertical displacement measurement by the P3-AT occurs when the unit is at rest. An accelerometer could however be integrated with the P3-AT to facilitate non-level surfaces found in actual pavement inspections

For accurately positioning, the authors have previously successfully integrated a virtual reference station

(VRS) system to the P3-AT to reach centimeter-level positioning accuracy (Chang et al., 2008). We are currently working towards the design and implementation of an inertial navigation system (INS) using an inertial measurement unit (IMU) and GPS with the P3-AT. The INS is capable of providing continuous estimates of P3-AT's position and orientation. The IMU coupled with the proper mathematical algorithms, is capable of detecting accelerations and angular velocities and then translating those to the current position and orientation of the P3-AT. The detailed pavement information (e.g. grade, cross fall, etc.) can be derived by the use of INS. The innovation of both systems can be used to improve pavement roughness inspections when used in conjunction with the P3-AT.

Acknowledgements

The authors would like to thank the National Science Council of Taiwan for their financial support provided under the project of NSC 96-2628-E-159-002-MY3, and the technical support received from the Keyence Corporation in Taiwan.

References

[1] Baus, R., and Hong, W. (1999). "Investigation and Evaluation of Roadway Rideability Equipment and Specifications." Federal Highway Administration/South Carolina Department of Transportation Report FHWA-SC-99-05, Columbia, South Carolina.

[2] Chang, J. R., Kang, S. C., Liu, M., Hsieh, S. H., Huang, T. C., and Lin, P. H. (2008). "An Autonomous Robot Equipped with the GPS Virtual Reference Station (VRS) System to Perform Pavement Inspections." International Symposium on Automation and Robotics in Construction (ISARC 2008), Vilnius, Lithuania, 141-147.

[3] FHWA-LTPP Technical Support Services Contractor. (2004) "2003 Comparison Testing of LTPP Profilers: Final Report." Long-Term Pavement Performance Team, Federal Highway Administration, U.S. Department of Transportation.

[4] Kelly, L. S., Leslie, T. G., Lynn, D. E. (2002). "Pavement Smoothness Index Relationships: Final Report." Publication No. FHWA-RD-02-057, U.S. Department of Transportation, Federal Highway Administration, Research, Development, and Technology, Turner-Fairbank Highway Research Center.

[5] Keyence Corporation. (2008). "LK-G Series Sensor Head Variations." <http://www.measurecentral.com/products/lkg/specifications1.php> (November 18, 2008)

[6] Ksaibati, K., McNamera, R., Miley, W., and Armaghani, J. (1999). "Pavement Roughness Data Collection and Utilization." Transportation Research Record 1655, Transportation Research Board, Washington, D.C.

[7] MobileRobots Inc. (2008). "The high performance all-terrain robot - P3-AT." <http://www.activrobots.com/ROBOTS/p2at.html> (October 21, 2008)

[8] Mondal, A., Hand, A. J., and Ward, D. R. (2000). "Evaluation of Lightweight Non-contact Profilers." Joint Transportation Research Program, Project Number: C-36-630, Purdue University, U.S. Department of Transportation, Federal Highway Administration.

[9] Perera, R. W., and Kohn, S. D. (2002). "Issues in Pavement Smoothness: A Summary Report." NCHRP Project 20-51.

[10] Sayers, M. W., and Karamihas, S. M. (1996). "Interpretation of Road Roughness Profile Data." Federal Highway Administration, FHWA/RD-96/101.

[11] The Transtec Group, Inc. (2008). "ProVAL-Intro-English." <http://www.roadprofile.com/data/proval/download/ProVAL-Intro-English.pdf> (October 27, 2008)

[12] Wang, H. (2006). "Road Profiler Performance Evaluation and Accuracy Criteria Analysis." Master thesis, Civil and Environmental Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Design of second-order sliding mode controllers for MR damper-embedded smart structures

H. Dalvand, H. T. Nguyen and Q. P. Ha

Faculty of Engineering and Information Technology, University of Technology Sydney, Broadway, NSW 2007, Australia, E-mail: {hdalvand, htn, quangha}@eng.uts.edu.au

Abstract

This paper presents the design of second-order sliding mode controllers for semi-active control using magneto-rheological (MR) dampers. The approach can be useful in applications involving shock absorbers but here our main concern is the suppression of building vibrations induced by dynamic loadings such as earthquakes or strong winds. The MR dampers have been of increasing interest in structural control as they are inexpensive to manufacture and have attractive properties such as small energy requirements, reliability and stability in operations, as well as a fast response of milliseconds. Challenges of MR damper structural control rest with the system's high nonlinearity due to the force-velocity hysteresis, and the constraint of the magnetisation current, required to be between its zero and maximal values. A variety of control algorithms have been applied, including the decentralized bangbang control, modulated homogeneous friction algorithm, clipped optimal control, Lyapunov-based control, and also non model-based intelligent schemes. In these techniques, the currents are usually obtained from the damping force indirectly rather than directly from the controller output. For direct current control, in this paper we propose second-order sliding mode controllers, which can satisfy the control constraint, provide high accuracy, retain robustness and remove chattering. The effectiveness of the proposed direct current control technique is verified, in simulations, on a benchmark building model subject to excitation of various scaled earthquake records.

Introduction

Control devices and methodologies for suppression of high-rise building vibrations caused by a dynamic loading source can be classified as passive dampers requiring no input power to operate, active dampers requiring a great deal of power to generate counteracting forces, and semi-active combining features of passive and active damping (Datta, 2003; Symans & Constaninou, 1999; Yoshida *et al.*, 2004).

In structural control, active control devices require a certain amount of energy to drive the actuators to accomplish the control objective. On the other hand, semi-active control needs a relatively small amount of driving power and the actuators can also be operated in the passive mode. The philosophy adopted in these approaches is to effectively absorb the vibration energy by modifying the control device physical characteristics.

For semi-active structural control, the use of magneto-rheological (MR) dampers has been of increasing interest in smart civil structures as they are inexpensive to manufacture, have reliable, stable and fail-safe operations, small energy requirements, and a fast response of milliseconds.

Given the advantages of MR dampers and semi-active control strategies, a number of controller designs have been proposed for the building control problem. In most of MR damper controllers developed so far, the current supplied to the dampers is quite often derived, from the required damping force obtained as the control signal, via a secondary current-control loop. In this paper, the direct current control approach for MR-dampers is proposed using second-order sliding mode (SOSM) controllers. The idea is to control directly the magnetisation current of the semi-active device in order to drive to zero not only the sliding function of the state variables but also higher-order time derivatives of the sliding function. The SOSM approach retains strong robustness of the system in the sliding mode, at the same time removes the chattering effect, provides even higher accuracy in realisation, and is suitable for control signals subject to constraints. These features make it ideal for direct current control of the MR damper used in the smart structures.

The remainder of the paper is organized as follows. The system description and the design of the proposed SOSM controller are included in Section 2. Simulation results are given in Section 3 to verify the effectiveness of the proposed approach. Finally, a conclusion is drawn in Section 4.

Control Design

Consider pairs of MR dampers, placed in a differential configuration on the 1st ,... k^{th} ,... and n^{th} floors of a building, with the control current vector $\mathbf{i} = [i_1 \cdots i_k \cdots i_n]^{\mathbf{T}}$ whose entries are constrained between zero and the maximal values. By defining the system state $\mathbf{y} = [\mathbf{x}^{\mathbf{T}} \quad \dot{\mathbf{x}}^{\mathbf{T}}]^{\mathbf{T}} \in \mathbb{R}^{2n}$, the state-space equation for the smart structure can be written as (Ha *et al.*, 2007):

$$\dot{\mathbf{y}} = \mathbf{A}\mathbf{y} + \mathbf{B}(\mathbf{y})\mathbf{i} + \mathbf{E} \quad , \tag{1}$$

in which

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{M}^{-1}\overline{\mathbf{K}} & -\mathbf{M}^{-1}\overline{\mathbf{C}} \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} \mathbf{0}_n & \mathbf{0}_n & \mathbf{0}_n \\ -m_1^{-1}(\overline{c}_{1_1}\dot{x}_1 + \overline{k}_{1_1}x_1) & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & -m_k^{-1}(\overline{c}_{1_k}\dot{x}_k + \overline{k}_{1_k}x_k) & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & -m_n^{-1}(\overline{c}_{1n}\dot{x}_n + \overline{k}_{1_n}x_n) \end{bmatrix}$$

$$\mathbf{E} = \mathbf{E_1} + \mathbf{E_2}, \qquad \mathbf{E_1} = \begin{bmatrix} \mathbf{0}_n \\ -m_1^{-1}\overline{\alpha}_1 z_1 \\ \vdots \\ -m_k^{-1}\overline{\alpha}_k z_k \\ \vdots \\ -m_n^{-1}\overline{\alpha}_n z_n \end{bmatrix}, \qquad \mathbf{E_2} = \begin{bmatrix} \mathbf{0} \\ \mathbf{\Phi} \end{bmatrix}\ddot{x}_g,$$

where $\mathbf{A}$ is the system matrix, $\mathbf{B}$ is the gain matrix and $\mathbf{E}$ is the disturbances (earthquake excitation and model uncertainties) of appropriate dimensions with notation given in (Kwok et al., 2006).

To design a structural controller that can perform satisfactorily in the presence of disturbances and uncertainty, different approaches have been proposed such as the linear-quadratic-Gaussian (LQG) control, sliding mode control (SMC), or Lyapunov-based control. However, smart structures embedded with MR dampers require the control currents to be constrained between zero and maximal magnetisation value, which normally results in some quantisation scheme, and hence, would affect the system performance. In this regard, it is attractive to use higher-order sliding mode controllers (Levant & Alelishvili, 2007) as they allow for using rates of change of the current as the control signal while having the ability to remove chattering and also to retain a wide range of robustness.

With differential dampers installed on the k^{th} floor, the motion equation for this floor is as below:

$$m_k\ddot{x}_k + c_k\dot{x}_k + k_k x_k = -[\overline{c}_{d1_k}\dot{x}_k + \overline{k}_{d1k}x_k + (\overline{c}_{d2}\dot{x}_k + \overline{k}_{d2}x_k)i_k + \overline{\alpha}_k z_{dk}] + m_k\ddot{x}_g \tag{2}$$

where $\dot{x}_k$ and $\ddot{x}_k$ are respectively the storey velocity and acceleration, i_k is the current supplied to the pair of dampers, and where $\overline{\alpha}_k = \overline{\alpha}_{10} + \overline{\alpha}_{11}i_k + \overline{\alpha}_{12}i_k^2$, $z_{d_k} = \tanh(\beta\dot{x}_k + \delta_k \operatorname{sign}(x_k)),$ and $\delta_k = \delta_{k0} + \delta_{k1}i_k$ (Ha et al., 2007). From (2), we obtain:

$$\ddot{x}_k = -m_k^{-1}\overline{c}_k\dot{x}_k - m_k^{-1}\overline{k}_k x_k - m_k^{-1}(\overline{c}_{d2}\dot{x}_k + \overline{k}_{d2}x_k)i_k - m_k^{-1}\overline{\alpha}_k z_{d_k} + \ddot{x}_g \ , \tag{3}$$

which has the general form of:

$$\ddot{x}_k = H_k(x_k, t) + G_k(x_k, t, i_k) \ . \tag{4}$$

For this dynamic equation, let us define a sliding function

$$\sigma_k = \dot{x}_k + \lambda_k x_k,\ \ \lambda_k > 0 \tag{5}$$

with time derivatives $\dot{\sigma}_k = \ddot{x}_k + \lambda_k \dot{x}_k$ and $\ddot{\sigma}_k = \dddot{x}_k + \lambda_k \ddot{x}_k$. Hence,

$$\ddot{\sigma}_k = \frac{dH_k}{dt} + \frac{dG_k}{dt} + \lambda_k (H_k + G_k)\,, \tag{6}$$

where

$$\frac{dH_k}{dt} = -m_k^{-1}\bar{c}_k \ddot{x}_k - m_k^{-1}\bar{k}_k \dot{x}_k + \dddot{x}_g$$

$$\frac{dG_k}{dt} = -m_k^{-1}\,\mathrm{sech}^2(\beta \dot{x}_k + \delta_k\,\mathrm{sign}(x_k))\bar{\alpha}_k \ddot{x}_k - m_k^{-1}(\bar{c}_{d2}\ddot{x}_k + \bar{k}_{d2}\dot{x}_k)i_k$$

$$-m_k^{-1}\left[\bar{c}_{d2}\dot{x}_k + \bar{k}_{d2}x_k + (\bar{\alpha}_{11} + 2\bar{\alpha}_{12}i_k)z_{d_k} + \bar{\alpha}_k\,\mathrm{sech}^2(\beta \dot{x}_k + \delta_k\,\mathrm{sign}(x_k)).\delta_{k1}\,\mathrm{sign}(x_k)\right]\frac{di_k}{dt}.$$

Therefore, by denoting $u_k = di_k / dt$, we can obtain the form:

$$\ddot{\sigma}_k = h_k(t, x_k, i_k) + g_k(t, x_k, i_k)u_k,\ \ h_k = \ddot{\sigma}_k\big|_{u_k=0},\ \ g_k = \frac{\delta}{\delta u_k}\ddot{\sigma}_k \neq 0 \tag{7}$$

where $h_k(t, x_k, i_k) = \dfrac{dH_k}{dt} + \lambda_k(H_k + G_k) - m_k^{-1}\,\mathrm{sech}^2(\beta \dot{x}_k + \delta_k\,\mathrm{sign}(x_k))\bar{\alpha}_k \ddot{x}_k - m_k^{-1}(\bar{c}_{d2}\ddot{x}_k + \bar{k}_{d2}\dot{x}_k)i_k$

$g_k(t, x_k, i_k) = -m_k^{-1}[\bar{c}_{d2}\dot{x}_k + \bar{k}_{d2}x_k + (\bar{\alpha}_{11} + 2\bar{\alpha}_{12}i_k)z_{d_k} + \bar{\alpha}_k\,\mathrm{sech}^2(\beta \dot{x}_k + \delta_k\,\mathrm{sign}(x_k)).\delta_{k1}\,\mathrm{sign}(x_k)]$.

Now, if we impose two conditions:

$$0 < k_{mk} \le g_k(t, x_k, i_k) \le k_{M_k},\ \ \text{and}\ \ |h_k(t, x_k, i_k)| \le C_k\,, \tag{8}$$

then according to (Levant, 2007), there exists a SOSM controller for $u_k(t)$ to drive σ_k and $\dot{\sigma}_k$ asymptotically to zero.

Assume now that (7) holds globally. Then (7) and (8) imply the differential inclusion

$$\ddot{\sigma}_k \in [-C_k, C_k] + [K_{m_k}, K_{M_k}]u_k\,, \tag{9}$$

where C_k, K_{m_k} and K_{M_k} are constants depending on the damper-embedded structure parameters defined in (7). Most SOSM controllers, for example (Levant, 2007; Polyakov & Poznyak, 2008; Levant & Pavlov, 2008, and Boiko et al., 2007), may be considered to steer $\sigma_k, \dot{\sigma}_k$ to 0 in finite time, which is essential for mitigation of quake-induced vibrations in structural control. Since inclusion (9) is not explicitly related to system (3), such controllers are obviously robust with respect to any perturbations, preserving (7). Hence, the problem is now to find a feedback control

$$u_k = \varphi_k(\sigma_k, \dot{\sigma}_k)\,, \tag{10}$$

such that all the trajectories of (9), (10) converge in finite time to the origin $\sigma_k = \dot{\sigma}_k = 0$ of the phase plane.

Differential inclusions (9), (10) are understood here in the Filippov sense (Filippov, 1988), which means that the right-hand set is enlarged in certain convexity and semi-continuity conditions. The function φ_k is assumed to be a locally bounded Borel-measurable function, which is physically true due to inertia of the magneto-rheological fluid. Indeed, in the smart structure control system, it represents the time rate of change of the magnetisation current to the MR dampers. A solution can therefore take any absolutely continuous vector function $(\sigma_k(t), \dot{\sigma}_k(t))$ satisfying (9), (10) for almost all t.

Design of SOSM controllers is greatly facilitated in the 2-dimensional phase plane with coordinates $\sigma_k, \dot{\sigma}_k$ by the simple geometry of any smooth curve that locally divides the plane into two regions. A number of known SOSM controllers may be considered as particular cases of a generalized 2-sliding homogeneous controller (Levant & Pavlov, 2008):

$$u_k = -r_{1_k} sign(\mu_{1_k}\dot{\sigma}_k + \lambda_{1_k}|\sigma_k|^{1/2} sign\sigma_k) - r_{2_k} sign(\mu_{2_k}\dot{\sigma}_k + \lambda_{2_k}|\sigma_k|^{1/2} sign\sigma_k),\ r_{1k}, r_{2k} > 0 \quad (11)$$

Drawing the two switching lines $\mu_{ik}\dot{\sigma}_k + \lambda_{ik}|\sigma_k|^{1/2} sign\sigma_k = 0,\ \mu_{ik}, \lambda_{ik} \geq 0,\ i = 1,2,\ \mu_{ik}^2 + \lambda_{ik}^2 > 0,$ $\mu_{1k}^2 + \mu_{2k}^2 > 0,\ \lambda_{1k}^2 + \lambda_{2k}^2 > 0$, in the phase plane, and considering various possible cases, one can readily check that it is always possible to choose r_{1k}, r_{2k} such that controller (11) yields finite-time stable responses. Indeed, if, for example, $\mu_{1k}, \lambda_{1k} > 0,$ then a 1-sliding mode can easily be induced on the line $\mu_{1k}\dot{\sigma}_k + \lambda_{1k}|\sigma_k|^{1/2} sign\sigma_k = 0$. If for each i one of the coefficients is zero, the twisting controller

$$u_k = -r_{1_k} sign(\sigma_k) - r_{2_k} sign(\dot{\sigma}_k) \quad (12)$$

is obtained with its convergence condition (Polyakov & Poznyak, 2008):

$$(r_{1k} + r_{2k})K_{mk} - C_k > (r_{1k} - r_{2k})K_{Mk} + C_k,\ (r_{1k} - r_{2k})K_{mk} > C_k. \quad (13)$$

Controller (11) may be considered as a generalization of the twisting controller, when the switching takes place on parabolas $\mu_{1k}\dot{\sigma}_k + \lambda_{1k}|\sigma_k|^{1/2} sign\sigma_k = 0$ instead of the coordinate axes.

An important class of SOSM controllers comprises the so-called quasi-continuous controllers, featuring control continuous everywhere except the SOSM $\sigma_k = \dot{\sigma}_k = 0$ itself. Since the 2-sliding condition is of dimension 2, the trajectory in general never hits the 2-sliding manifold. Hence, the control signal, or the time derivative of the damper magnetisation current in (7), remains a time-continuous function all the time. As a result, chattering is significantly reduced. In this paper, we select the following SOSM controller from such a family, as given in (Levant, 2007):

$$u_k = -\alpha_k \frac{\dot{\sigma}_k + \beta_k|\sigma_k|^{1/2} sign\sigma_k}{|\dot{\sigma}_k| + \beta_k|\sigma_k|^{1/2}}. \quad (14)$$

This controller is continuous everywhere except of the origin and vanishes on the parabola $\dot{\sigma}_k + \beta_k|\sigma_k|^{1/2} sign\sigma_k = 0$. With sufficiently large α_k there are such numbers ρ_{1k}, ρ_{2k}, where $0 < \rho_{1k} < \beta_k < \rho_{2k}$, that all the trajectories enter the region between the curves $\dot{\sigma}_k + \rho_{1_k}|\sigma_k|^{1/2} sign\sigma_k = 0$ and remain there.

As described in (7), since the SOSM control is the derivative of the damper current, the current itself is obtained by integration:

$$i_k = \int u_k dt = \int -\alpha_k \frac{\dot{\sigma}_k + \beta_k|\sigma_k|^{1/2} sign\sigma_k}{|\dot{\sigma}_k| + \beta_k|\sigma_k|^{1/2}} dt. \quad (15)$$

Simulation Results

For illustration, a 3-storey structure is considered in which differential dampers are placed on the first floor. A block diagram of the structure is depicted in Fig. 1. The parameters for the smart structure are as below

$$\mathbf{M} = \begin{bmatrix} 98.3 & 0 & 0 \\ 0 & 98.3 & 0 \\ 0 & 0 & 98.3 \end{bmatrix}(kg),\ \mathbf{C} = \begin{bmatrix} 175 & -50 & 0 \\ -50 & 100 & -50 \\ 0 & -50 & 50 \end{bmatrix}(Ns/m),$$

$$\mathbf{K} = \begin{bmatrix} 12.00 & -6.84 & 0 \\ -6.84 & 13.80 & -6.84 \\ 0 & -6.84 & 6.84 \end{bmatrix} \times 10^5 (N/m)$$

$$i_1 = i,\ \ i_2 = i_3 = 0\ ,\ \overline{\alpha}_1 = 64 + 1836i - 488i^2,\ \ \overline{\alpha}_2 = \overline{\alpha}_3 = 0\ ,$$

$$z_{d1} = \tanh(100\dot{x}_1 + (0.58 + 0.30i)\,\mathrm{sign}(x_1)),\ \ z_{d2} = z_{d3} = 0\ .$$

A. Controlled responses

Firstly, a number of excitations including step, random, sinusoidal and square waveforms are considered. The responses of the structure, for example to random and sinusoidal excitations are illustrated in Figs. 2 and 3, respectively. According to the figures, the proposed controller can mitigate effectively the affect of the external disturbances by directly controlling the damper current.

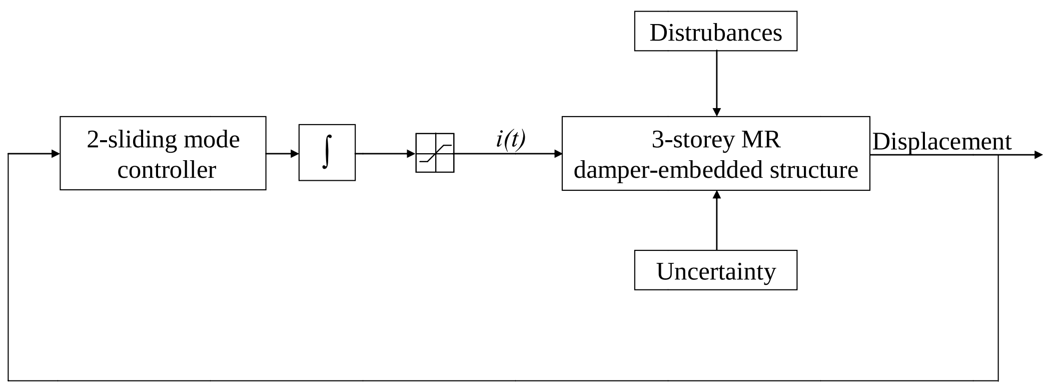


Figure 1. Second order sliding mode controlled 3-storey MR damper-embedded structure

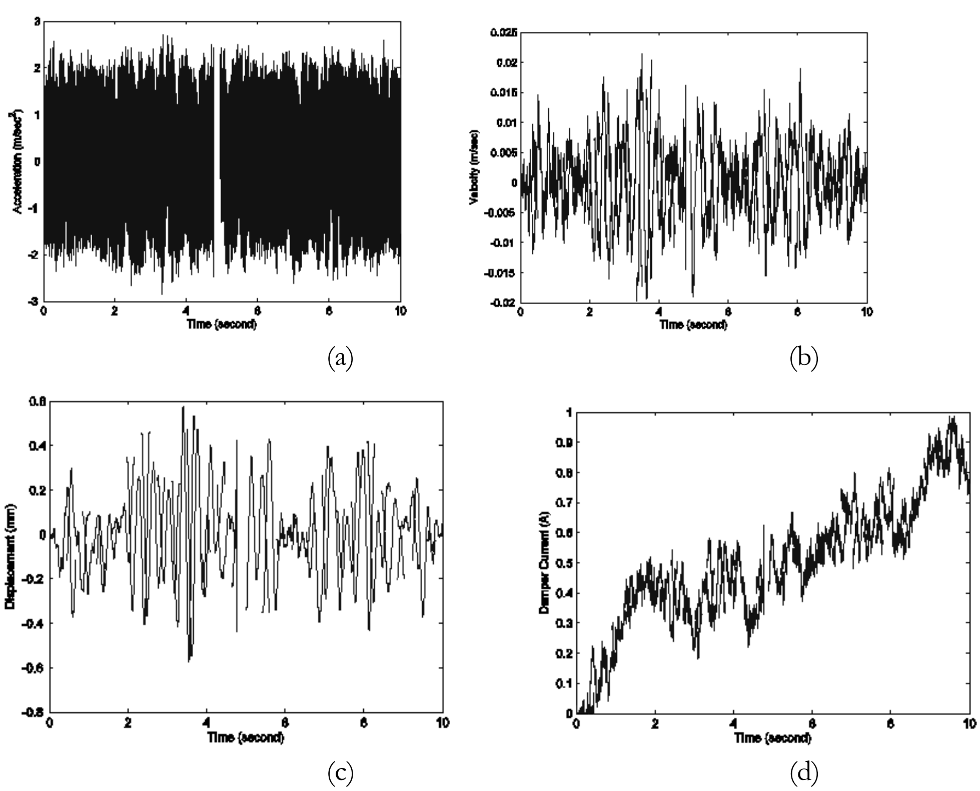
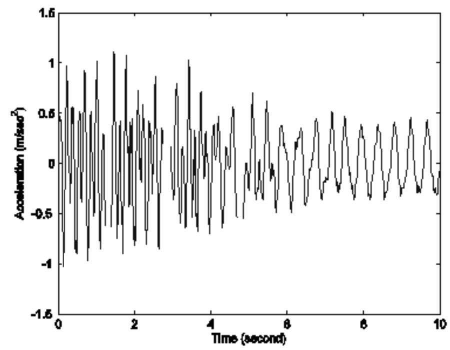
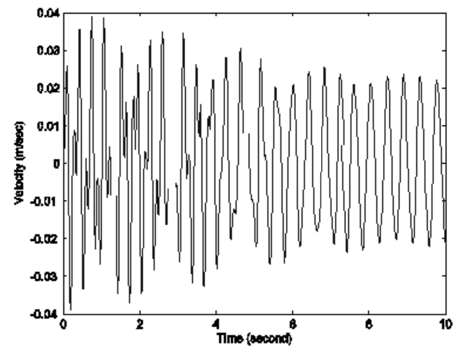
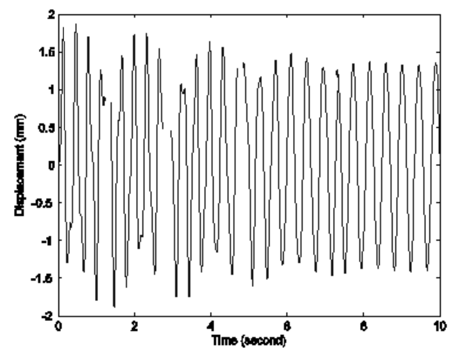


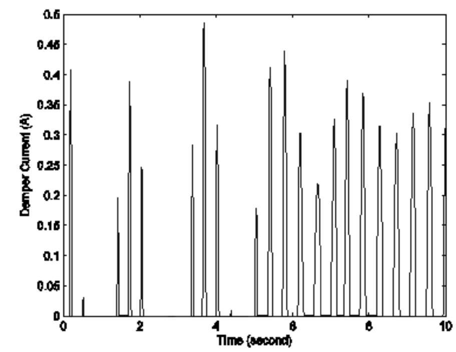
Figure 2. Random excitation responses: (a) 1st floor Acceleration, (b) 1st floor Velocity, (c) 1st floor Displacement, and (d) Damper Current





(a) (b)





(c) (d)

Figure 3. Sinusoidal excitation responses: (a) 1st floor Acceleration, (b) 1st floor Velocity, (c) 1st floor Displacement, and (d) Damper Current.

We consider next responses to the scaled records of known earthquakes. The first-storey time responses for the Kobe earthquake are shown in Fig. 4. The responses display significant reductions in displacement, velocity and acceleration. Similar responses can be obtained for scaled records of the El-Centro, Northridge, and Hachinohe. The SOSM controller indicates the system stability in most of earthquake period except where the magnitude is too large. However, the derivative returns to negative and the building structure under control becomes stable.

Performance Evaluation

Apart from six performance criteria given in (Ha et al., 2007), we consider here with reference to (Spencer et al., 1999) further four evaluation criteria, two for peak responses and two for RMS responses. They are:

Peak inter-storey drift ratio

$$J_7 = \max\left\{\frac{\max\{|\bar{x}_{k,c}(t)|\}}{\max\{|x_{k,u}(t)|\}}\right\}, \qquad (16)$$

whereby the maximum drifts are normalized with respect to the uncontrolled peak displacement, subscript $k = 1,\ldots,3$ stands for the storey index and subscripts c, u denote controlled and uncontrolled cases.

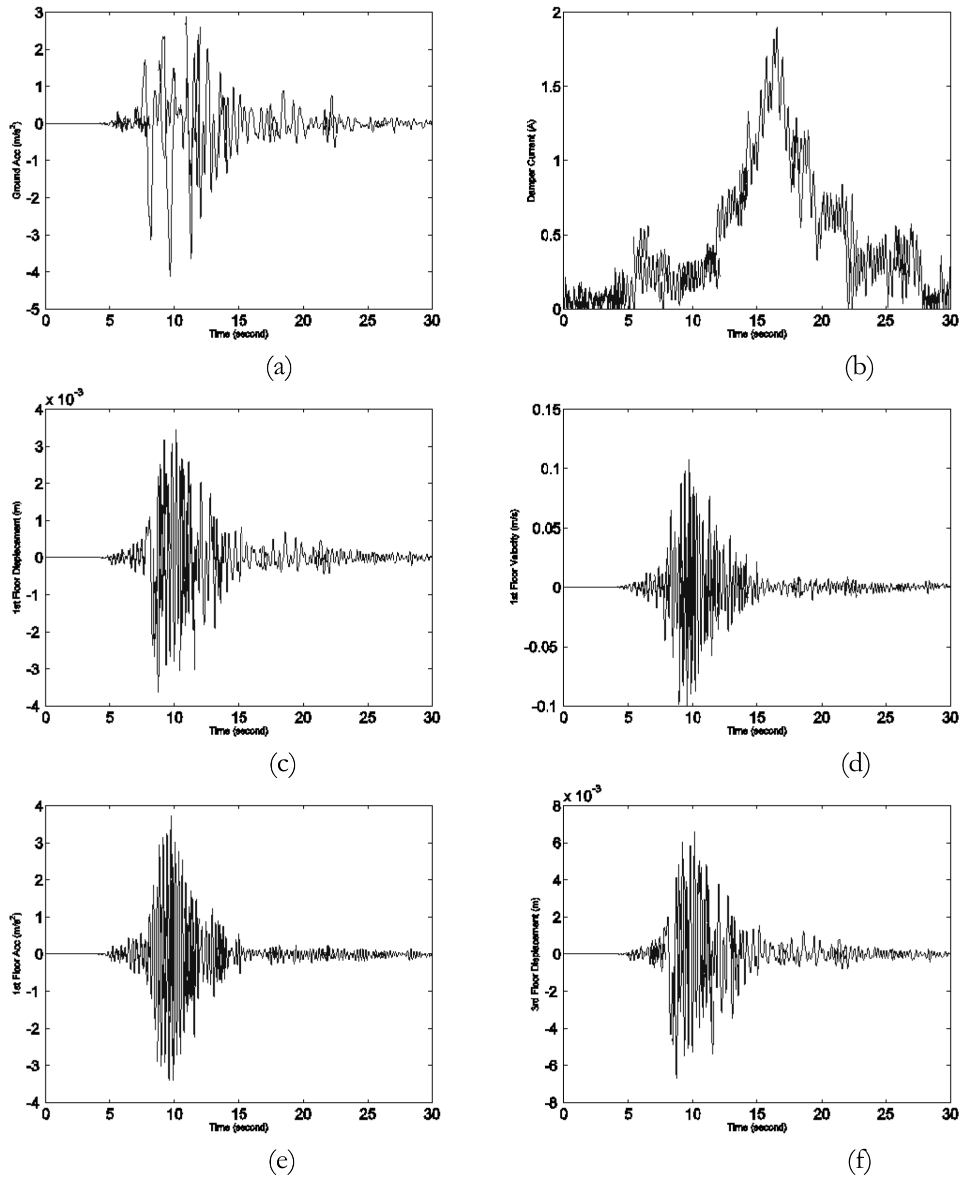


Figure 4. *Kobe* Earthquake Responses: (a) Excitation, (b) Damper Current, (c) 1st floor Displacement, (d) 1st floor Velocity, (e) 1st floor Acceleration, and (f) 3rd floor Displacement

Peak storey acceleration ratio

$$J_8 = \max\left\{ \frac{\max\{|\ddot{x}_{k,c}(t)|\}}{\max\{|\ddot{x}_{k,u}(t)|\}} \right\}, \tag{17}$$

whereby the accelerations are normalized by the peak uncontrolled acceleration.

Maximum RMS inter-storey drift ratio

$$J_9 = \max\left\{ \frac{\tilde{\overline{x}}_{k,c}(t)}{\tilde{\overline{x}}_{k,u}(t)} \right\}, \tag{18}$$

which evaluates the ability of minimizing the maximum RMS inter-storey drift due to all admissible ground motions. The notation $\tilde{x} = \sqrt{T^{-1}\sum\{\delta_t x_k^2(t)\}}$ is for the root-mean-square (RMS) values, δ_t is the sampling time, and T is the total excitation duration.

Maximum RMS storey acceleration ratio

$$J_{10} = \max\left\{\frac{\tilde{\ddot{x}}_{k,c}(t)}{\tilde{\ddot{x}}_{k,u}(t)}\right\} \tag{19}$$

that is given in terms of the maximum RMS absolute acceleration with respect to the uncontrolled case.

Table I below summarizes all the criteria evaluated using the simulated responses with the proposed SOSM controller, typically, for the second floor. As can be seen, all the corresponding ratios using the SOSM controller are much smaller than those obtained in the uncontrolled case and further improved from the Lyapunov-based method (Ha et al., 2007).

Table I. Response ratios:

2nd Floor	J_1	J_2	J_3	J_4	J_5	J_6	J_7	J_8	J_9	J_{10}
El-Centro	0.148	0.132	0.145	0.040	0.069	0.076	0.068	0.078	0.126	0.555
Kobe	0.152	0.127	0.146	0.043	0.0620	0.079	0.061	0.081	0.127	0.454
Hachinohe	0.103	0.061	0.099	0.026	0.041	0.051	0.042	0.055	0.059	0.334
Northridge	0.130	0.103	0.125	0.032	0.047	0.060	0.047	0.070	0.101	0.307

Conclusions

We have presented an effective scheme for semi-active control of smart structures embedded with pairs of MR dampers. Using the proposed control system, the building structures are shown to be capable of effectively suppressing vibrations due to earthquakes by directly controlling the damper magnetization currents. Differential configuration for the dampers is used to remove the problem arising from damper offset forces. For this semi-active structural control system, a second-order sliding mode controller is proposed to obtain the time rate of change of the supplied currents to the dampers. These magnetisation currents, after integration, can efficiently control the fluid to yield required damping forces for the structures with provident power consumption and improved control performance. Simulation results for a three-floor building model are evaluated using a set of performance criteria. The results obtained demonstrate the effectiveness of the proposed scheme under constraints of the control signals in mitigation of seismic vibrations of MR damper embedded smart structures.

Acknowledgement

This work is supported by the Australian Research Council (ARC) Discovery Project DP0559405, the UTS Centre for Built Infrastructure Research, and in part by the Centre of Excellence programme, funded by the ARC and the New South Wales State Government.

References

[1] Boiko, I., Fridman, L., Pisano, A., and Usai, E. (2007). "Analysis of chattering in systems with second-order sliding modes." IEEE Trans. on Automatic Control, 52(11), 2085-2102.

[2] Datta, T. K. (2003). "A state-of-the-art review on active control of structures." ISET J. Earthquake Technology, 40(1), 1-17.

[3] Filippov, A. F. (1988). Differential equations with discontinuous right-hand side. Kluwer Academic Publishers, Dordrecht.
[4] Ha, Q. P., Kwok, N. M., Nguyen, M. T., Li, J., and Samali, B. (2007). "Mitigation of seismic responses of building structures using MR dampers with Lyapunov-based control." J. Struct. Control and Health Monitoring, 15(4), 604-621.
[5] Kwok, N. M., Ha, Q. P., Nguyen, T. H., Li, J., and Samali, B. (2006). "A novel hysteretic model for magnetorheological fluid dampers and parameter identification using particle swarm optimization." Sensors & Actuators: A. Physical, 132(2), 441-451.
[6] Levant, A. and Alelishvili, L. (2007). "Integral high-order sliding modes." IEEE Trans. on Automatic Control, 52(7), 1278-1282.
[7] Levant, A. (2007). "Construction principles of 2-sliding mode design." Automatica, 43(4), 576-586.
[8] Levant, A. and Pavlov, Y. (2008). "Generalized homogeneous quasi-continuous controllers." Int. J. Robust and Nonlinear Control, 18(4-5), 385-398.
[9] Polyakov, A. and Poznyak, A. (2008). "Lyapunov function design for finite-time convergence analysis of twisting and super-twisting second order sliding mode controllers." Proc. of the 10th Int. Workshop on Variable Structure Systems (VSS'08), 8-10 Jun., Antalya, Turkey, 153-158.
[10] Spencer, B. F., Dyke, S. J. and Deoskar, H. S. (1999). "Benchmark problems in structural control, Part I: Active mass driver system." J. Earthquake Eng. & Struct. Dynamics, 27(11), 1127-1139.
[11] Symans, M. D. and Constaninou, M. C. (1999). "Semi-active control systems for seismic protection of structures: A state-of-the-art review." J. Eng. Struct., 21(6), 469-487.
[12] Yoshida, O. and Dyke, S.J. (2004). "Seismic control of a nonlinear benchmark building using smart dampers." J. Eng. Mechanics, 130(4), 386–392.

Motion Control of Excavator with Tele-Operated System

Dongnam Kim[1], Kyeong Won Oh[2], Daehie Hong[3#], Yoon Ki Kim[4] and Suk-Hie Hong[5]

[1]Mechatronics & Field Robotics Laboratory, Division of Mechanical Engineering, Korea University, 5Ga-1, Anam-Dong, Sungbuk-Gu, Seoul 136-701, Korea; PH +82-2-3290-3765; FAX +82-2-3290-3864;
e-mail : smileast@korea.ac.kr
[2]Mechatronics & Field Robotics Laboratory, Division of Mechanical Engineering, Korea University, 5Ga-1, Anam-Dong, Sungbuk-Gu, Seoul 136-701, Korea; PH +82-2-3290-3765; FAX +82-2-3290-3864;
e-mail : locoski@korea.ac.kr
[3]Division of Mechanical Engineering, Korea University, 5Ga-1, Anam-Dong, Sungbuk-Gu, Seoul 136-701, Korea; PH +82-2-3290-3369; FAX +82-2-926-9290;
e-mail : dhhong@korea.ac.kr
[#]corresponding author
[4]Naekyung Engineering Co., Ltd., 5-1, Sewon building, Yangjae-Dong, Seocho-Gu, Seoul 137-130, Korea; PH +82-2-3461-2900; FAX +82-2-574-4958;
e-mail : nk2900@empal.com
[5]Naekyung Engineering Co., Ltd., 5-1, Sewon building, Yangjae-Dong, Seocho-Gu, Seoul 137-130, Korea; PH +82-2-3461-2900; FAX +82-2-574-4958;
e-mail : hongcon@korea.com

Abstract

The number of dismantling sites for buildings has been rapidly increased and these processes are very dangerous. So many research papers have been published addressing the development of a remote controller for the dismantling equipments. In this paper, a novel concept of applying tele-operated device is proposed for the remote control of excavator-like dismantling equipment. As a tele-operated system, this controller is designed to improve the operability of the excavator. This paper also includes all the necessary kinematic analysis to design the tele-operated system. In order to explore the feasibility of the device, modeling of the tele-operated system is used to demonstrate its convenience. Then, the hydraulic system of excavator which uses the proportional valve system is also studied. And basic motion control simulations to the real excavator working at construction site are conducted with designed tele-operated system. And operator can control the real excavator intuitively with this new model of tele-operated system.

Keywords: Haptic Device, Excavator, Inverse Kinematics, Dismantling Process, Hydraulic System

Notations

Z_{e1} : Axis of swing joint
Z_{e2} : Axis of boom joint
Z_{e3} : Axis of arm joint
Z_{e4} : Axis of bucket joint
a_2 : Distance between Z_{e2} and Z_{e3} (Length of boom)
a_3 : Distance between Z_{e3} and Z_{e4} (Length of arm)
r_{24} : Distance between Z_{e2} and Z_{e4}
θ_{ei} : Rotational angle of each link of excavator (i=1, 2, 3, 4)

1. Introduction

The construction technology has been developed for thousands of years, so the latest buildings are about several hundred meters high. Furthermore, how tall buildings can be stably constructed has been a main theme in architecture. Recently, however, how to dismantle an old high-rise building has also been focused by many researchers. Dismantling methods have been also developed for centuries, and among these methods, mechanical dismantling has been very common. For example, the excavator which is used to construct a new designed building is also utilized for the heavy equipment for dismantling processes. Especially crusher and breaker are the attachments for the excavator to crack the building to pieces. This excavator which is equipped with the attachment can be easily seen at dismantling sites. However dismantling processes are very dangerous, so many safety-related accidents happen in many sites. Furthermore, operator has to ride on the excavator and this makes the dismantling process more dangerous. Therefore, remote control device is necessary to guarantee an operator's safety.

There have been some recent researches related to the tele-operation of the excavator. First, kinematics analysis and simulation of construction devices were conducted by Frimpong and Stentz. Furthermore, Frankel studied remote control of excavator with commercial haptic device, Phantom. Finally, haptic device of excavator which was operated in underwater was studied by Hirabayashi and Sasaki. And last, Phantom device is also studied by Kim and Feygin. And another tele-operating system for excavator is conducted by Oh.

In this paper, tele-operated device for dismantling process is main subject. This device is designed based on the kinematics of the excavator, which can cover 3-dimensional workspace. This kinematics analysis of the excavator is studied. This paper also includes all the necessary kinematic analysis to design the tele-operated device. In order to explore the feasibility of the device, modeling of this device is used to demonstrate its convenience. Then, the hydraulic system of excavator which uses the proportional valve system is also studied. And last, basic motion control simulations to the real excavator with sensor feedback mechanism are conducted with designed tele-operated system. As a result, this device can be proposed to confirm the feasibility of the tele-operated system. And one hand control can be possible to move the real excavator with this new model.

2. Kinematics of Excavator

In the world, there are many kinds of backhoes. However, in this paper, excavator whose motions are called like swing, boom, arm, and bucket, is chosen. This excavator may be easily seen in many construction sites. Based on this motion, Figure 1 shows 4 coordinates of each links of the excavator. First, swing joint axis of excavator is represented by Z_{e1}, and the remaining three joint axes, boom, arm, bucket are also declared.

After the modeling of excavator, forward and inverse kinematics analysis is necessary. When the haptic devices are designed, the inverse kinematics analysis is more important. Because the operator wants to control the pose of last link, bucket which should be determined by the tele-operated system. With this reason, according to the bucket's pose, four joint angles of excavator are controlled by the inverse kinematics analysis. These joint angles can be calculated with following equations.

$$\theta_{e2} = \tan^{-1}\left(\frac{r_{24}(y)}{r_{24}(x)}\right) + \cos^{-1}\left(\frac{a_2^{\ 2} + r_{24}^{\ 2} - a_3^{\ 2}}{2a_2 r_{24}}\right) \tag{1}$$

$$\theta_{e3} = \pi + \cos^{-1}\left(\frac{a_2^{\ 2} + a_3^{\ 2} - r_{13}^{\ 2}}{2a_2 a_3}\right) \tag{2}$$

Degree of freedom of excavator is four, at the same time, bucket can move in 3-dimensional space. In other words, the bucket movement can be represented with haptic stylus in 3D space. However, when the real digging operations are examined, it is observed that the swing motion is nearly not used. That is, the main motion of excavator is controlled by 3 DOF without swing, which determines the bucket's 2D motion. The swing motion is usually used to move the dug earth to other side. If the swing motion is separated from the digging motion and can be controlled by the haptic device's another reserved joint, operator can control the excavator more easily.

Therefore, inverse kinematics analysis of excavator which is 2-dimensional kinematics problem to control the boom and arm is necessary. However, with this control method, 2D plane motion of bucket joint cannot be controlled intuitively. The position of bucket joint can be acquired with the two angles of boom and arm. So, if operator wants to control the only boom angle or arm angle, it is difficult to control each angle with this 2D plane control method. So, design concept of tele-operated system will be shown next chapter.

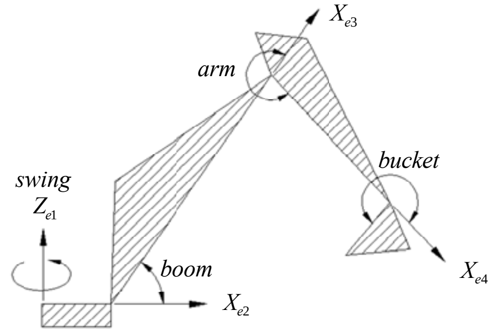


Figure 1 Kinematic model of excavator

3. Design of Tele-Operated System

In previous research, a haptic device was designed, such that operator can work with intuitive bucket control. So, novice operator can use the excavator easily unlike the conventional controller. The degree of freedom of haptic device is 4, because the excavator has 4 degrees of freedom. First, swing and bucket control is directly connected to two joints of haptic device. And 2-dimensional position control is mapped with the position of the bucket joint. But this method makes the control somewhat unintuitive. So new concept design is proposed like Figure 2 and Figure 3. This is not a fine design, but shows the concept design of tele-operated device. And Figure 4 is a last drawing of haptic device. Operator put ones elbow on the table of the tele-operation system. And grab the haptic stylus and move ones arm freely. This motion controls the full motion of the excavator. First, d_{h1} controls the boom motion. Second, θ_{h4} which represents the angle of wrist of operator controls the bucket motion directly. And the angle θ_{h3} between the lower arm and horizontal plane controls the arm angle of excavator. And last θ_{h2} controls the swing motion of excavator. With this design concept, operator can control the each angle directly.

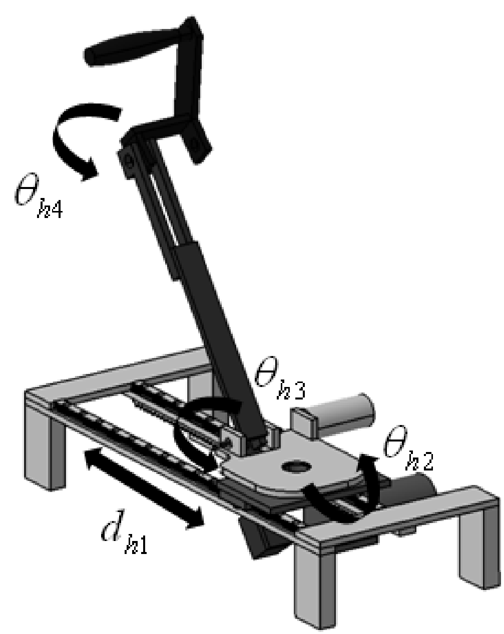
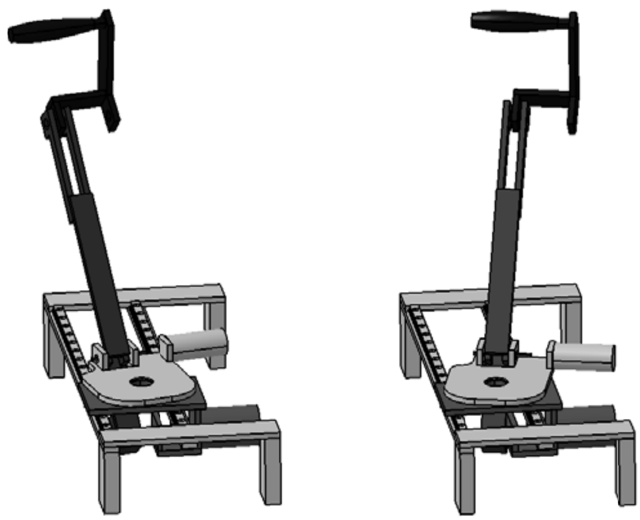


Figure 2 Design concept of tele-operated system



(a) Swing to the left direction (b) Swing to the right direction

Figure 3 Swing motion control of excavator with tele-operated system

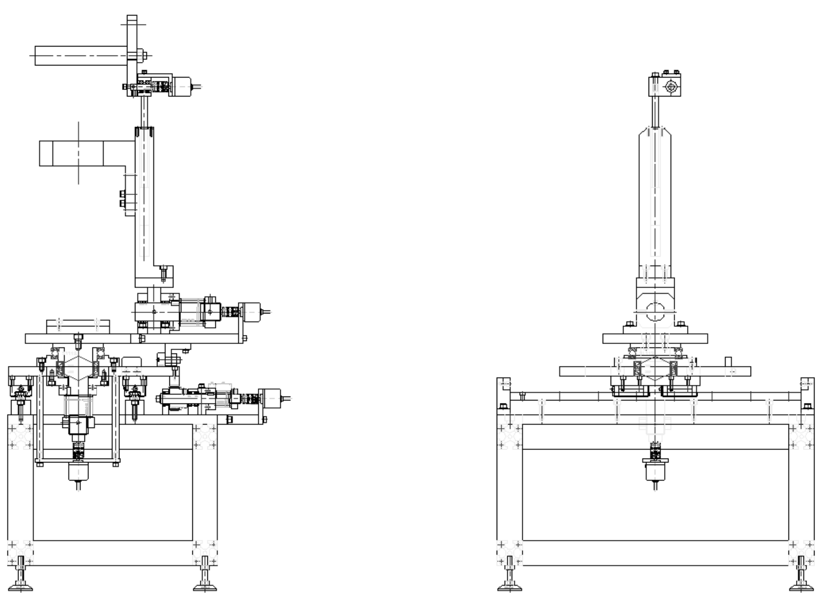


Figure 4 Final design of tele-operated system before manufacturing

4. Control of Excavator

There are some sensors to feedback control. Inclinometers and gyro sensor are used to feedback pose value of excavator. DAQ 6024E board (NI) is used to get the values of sensors. Industrial PC is also set up on the excavator and values of sensors can be monitored. This feedback mechanism circuit is shown in Figure 5.

Three inclinometers and one gyro sensor are necessary. Inclinometers are attached to get the angle value of each link, boom, arm and bucket. And gyro sensor is installed on the base frame of excavator to get the pose. The angles of links are acquired with Visual C++ program through DAQ board, and then these values are used to control the motion of the excavator. PI control result is shown in Figure 6. The reference inputs are 70 degree (boom) and 200 degree (arm). If controller has only P controller, there is steady state error which is about 5 degrees. But there is almost 1 degree of steady state error with PI controller. However friction of boom joint is larger than that of arm joint, control result of boom is not excellent in this paper.

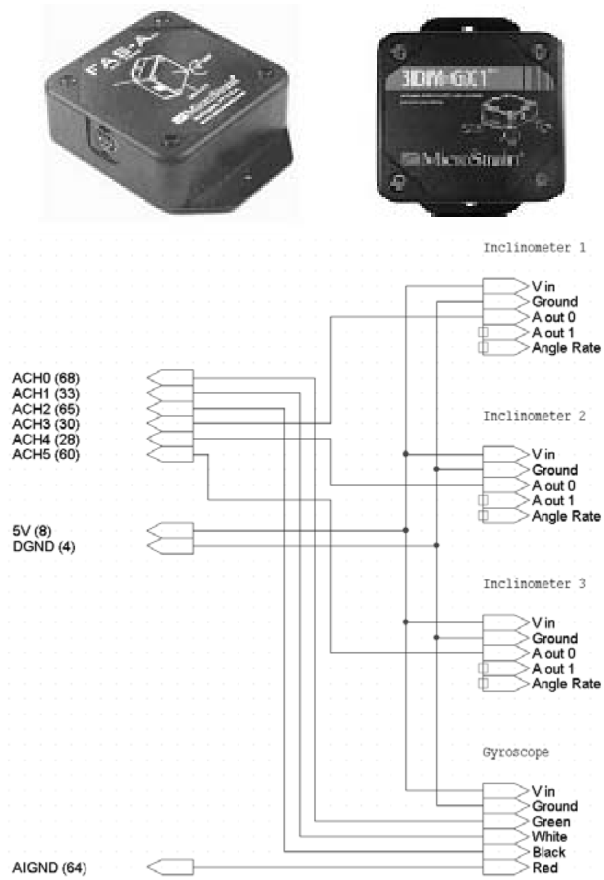
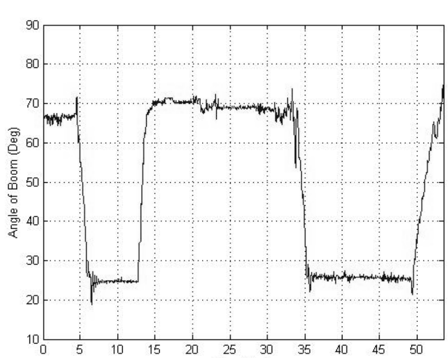
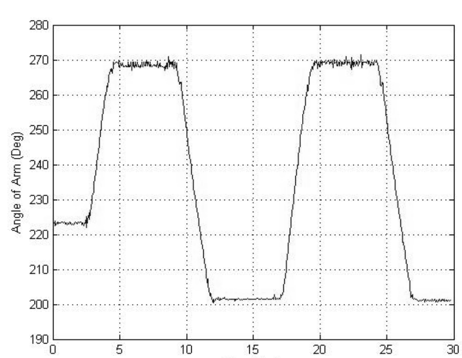


Figure 5 Picture of sensor and circuit of sensor feedback with DAQ board

For the control of excavator, hydraulic system should be studied first. Figure 7 is the picture of real excavator whose name is SOLAR 015. Main hydraulic control valve system has 9 degrees of freedom in Solar 015. So, there are 9 hydraulic ports in control system. And among these ports, first, fourth, fifth and eighth control valve is pilot valve and these control the main links of excavator (swing, boom, arm and bucket). Pilot valves are controlled with the joystick of excavator installed on the both side of operator seat. This is original model of excavator but it has to be remodeled with some equipment. First, proportional valve system which is shown in Figure 8 is adapted to control electrically. There are 8 proportional valves connected to the pilot valves of main hydraulic valve. So with these proportional valves, operator can control the excavator electrically from remote sites.



(a) Boom angle control result



(b) Arm angle control result

(Ref=70degree) (Ref=270degree)

Figure 6 Angle control result of excavator with PI controller

Figure 7 Picture of excavator which is SOLAR 015

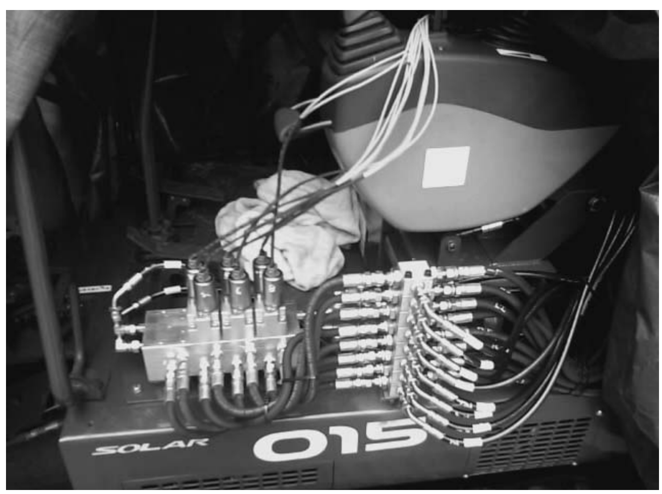


Figure 8 Proportional valve systems to control the excavator

Figure 9 Remote controller of excavator (joystick and receiver)

Basically, this proportional valve can be controlled with wireless joystick system, and this joystick system is shown in Figure 9. This is same as the joystick of excavator. And RF transmitter module is installed inside of this joystick and the excavator is equipped with RF receiver module. So operator can control the excavator with this wireless joystick system at the remote sites. And to use the original joystick of excavator, 8 shuttle valves are also equipped. Operator can use both the remote joystick and excavator joystick to operate the excavator.

In this paper, main object is to make the tele-operated system, so, another system should be added. First, electrical control unit should be adapted, and analog signal can be transferred to the proportional valve. DAQ 6024 board (National Instrument) is used to make the analog signal from 0V to 3V. When the input signal is 1.5V to the RF module, link of excavator doesn't move. If 3V is connected, link moves one direction, and 0V makes opposite direction movement. There are 2 analog output channels in DAQ 6024 board, boom and arm control valve is connected to this analog output channel. So, full control of excavator, 2 DAQ boards are necessary. In the future, more developed DAQ board should be installed to minimize the system. And control program is made with visual C++. And this program can get the value of inclinometer. It uses the input channel of DAQ board. And there is also PI controller to move the boom and arm to desired position. This is required to control the motion of excavator directly. Simulation results of this PI control was shown in Figure 6.

5. Conclusions

In this paper, new concept of tele-operated system was designed to control the excavator. First, kinematics of excavator was analyzed and hydraulic system of excavator is studied. Next sensor feedback mechanism is also included to get the angles of each link. Last, motion simulation is shown with the SOLAR 015 excavator. With this result, tele-operated system is proven to be very simple method. In the future, haptic device will be manufactured and the more intuitive and comfortable control of excavator will be researched.

Acknowledgement

This research was supported by a grant (code: 06 Construction Core B04) from Construction Core Technology Program funded by Ministry of Construction & Transportation of Korean government and BK21.

References

[1] Stentz, A., Bares, J., Singh, S. and Rowe, P. (1999). "A Robotic Excavator for Autonomous Truck Loading," Autonomous Robots, Vol. 7, No. 2, pp. 175-186.
[2] C¸avu¸so˘glu, M. C. and Feygin, D. (2001). "Kinematics and Dynamics of Phantom(TM) model 1.5 Haptic Interface," Sensable, pp. 1-13.
[3] Joseph, G. F. (2004). "Development of a Haptic Backhoe Testbed," A Thesis of Master of Science, School of Mechanical Engineering, Georgia Institute of Technology, pp. 14-30.
[4] Frimpong, S. and LI, Y. (2005). "Virtual prototype simulation of hydraulic shovel kinematics for spatial characterization in surface mining operations," Int. J. of Surface Mining, Reclamation and Environment, Vol. 19, No. 4, pp. 238-250.
[5] Doosan Infracore, Co., Ltd. (2005). "SOLAR 015 Parts Book".
[6] Hirabayashi, T. and Yamamoto, T. (2006). "Experiment of Teleoperation of Underwater Backhoe with Haptic Information," Proceedings 23rd Int. Sym. on Automation and Robotics in Construction, pp. 36-41.
[7] Sasaki, T., Nagai, T. and Kawashima, K. (2006). "Remote Control of Backhoe for Rescue Activities Using Pneumatic Robot System," Proceedings of the IEEE International Conference on Robotics and Automation, pp. 3177-3182.
[8] Dongnam, K., Kyeong Won, O., Daehie, H., Jong-Hyup, P. and Sukhie, H. (2007). "Design and Simulation of A Haptic Device for Excavator with Phantom 1.5," Asian Symposium for Precision Engineering and Nanotechnology.
[9] Kyeong Won, O., Dongnam, K., Daehie, H., Jong-Hyup, P. and Sukhie, H. (2007). "Design of a haptic device for excavator equipped with crusher," ISARC2007.

On-line Process Management of Pavement Laying Using Wireless Communication Technologies

Viljamaa E.[1], Kilpeläinen P.[1], Pentikäinen V.[1], Sarjanoja E-M.[1] and Heikkilä R.[2]

[1]VTT Technical Research Centre of Finland, P.O.BOX 1100 (Kaitoväylä 1), FIN-90571 Oulu, Finland, firstname.lastname@vtt.fi
[2]University of Oulu, Research Unit of Construction Technology, P.O.BOX 4400 (Kiilakiventie 1), FIN-90014 Oulu, Finland, firstname.lastname@oulu.fi

Abstract

The research goal has been synchronizing the process of asphalt production, asphalt transportation and pavement laying by connecting different machines together using wireless technologies. Sensors and GPS positioning units were installed to the machines for measuring process variables (e.g. mass temperature) and detecting events (e.g. loading, unloading). The first prototype system (autumn 2007) was based on embedded wireless measurement and data collection modules. A more recent system uses a vehicle PC and a separate embedded system module with sensor interfaces and a short-range radio transmitter. Vehicles communicate with the database through a GPRS link and use a short-range radio as an RFID for authentication and a direct communication channel with the other vehicles. GPS position information is used for context recognition together with the short-range radio and sensor information. Process information is continuously stored to a database, to be used for quality control and process optimization.

Keywords: Automatic data collection, pavement laying, process management, wireless communication

1. Introduction

Since recent road pavement processes are not optimal and many process variables are usually unknown this research focused on the process management and the quality data collection of the asphalt paving process [Peyret et al. 2000]. A system for monitoring and controlling asphalt pavement process was developed by connecting the machines in this process together using real-time wireless communication in a similar way that production machines are typically connected inside a production plant. This enables information sharing between the machine operators and automatic data collection using sensors installed to the machines [Kilpeläinen, Heikkilä & Parkkila 2007]. This concept was evaluated by building two prototype systems that were tested at construction sites. The conceptual approach was chosen for reasons that firstly, most of the earlier research of road construction concerned the subprocesses like vehicle tracking [Lu et al. 2007], compacting [Peyret et al. 2000] or simulation [Nassar, Thabet & Beliveau 2003] and there were no automated processes researched and secondly, recent systems in use are mainly manual and not automated at all. Both of these prototype systems had GPS positioning unit and a wireless communication link (GPRS) that were installed to the machines in order to collect data from the paving process and to save this data to the database. These actions were aiming to the ultimate goal that is smoothly and continuously streaming paving process where resources are used with optimal efficiency.

One of the prototype systems was based on embedded measurement modules. The goal was to collect data as automatically as possible with minimum user interaction needed. This system was tested autumn 2007. [Kilpeläinen, Heikkilä & Parkkila 2007]

The more recent prototype system mainly described in this paper had a main focus on machine interaction and designing user interfaces for different machine operators were developed. Another new feature was a direct short-range communication link between the machines. Some more advanced features, like automatic reporting and billing, were also tested. In this prototype the work machines were equipped with in-vehicle PCs and embedded measurement and radio modules. The second prototype was tested autumn in 2008.

2. Prototype System Overview

2.1 Prototype system overview

The prototype system included three different operators: asphalt mixing plant unit in the asphalt mixing station and a vehicle unit with variations to a truck and to a paver. A compactor that is a natural part of paving process was left out of the scope in this time, since it is not having that many interactions between other operators. Both of the unit types had two different communication methods: long-range and short-range techniques. Also the external server with database and web pages was supporting the functionalities of the units. Figure 2 illustrates the complete prototype system.

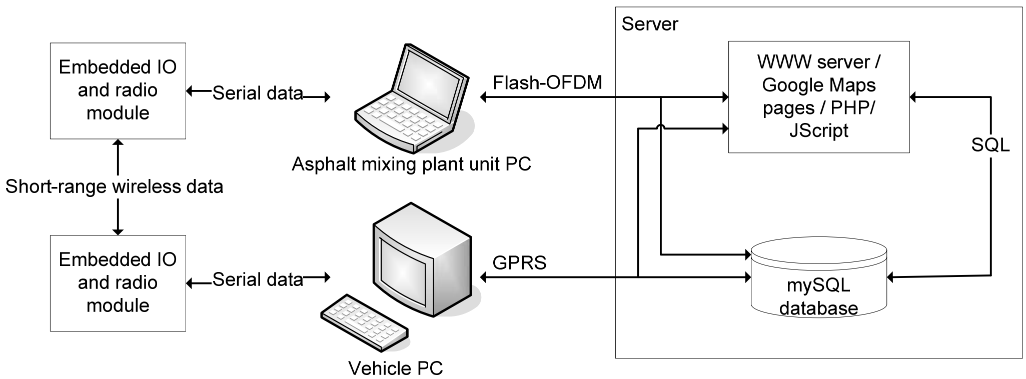


Figure 2. The prototype system.

2.2 Paver and Truck units

The paver and the truck units of the system consisted a SUNIT D12 vehicle PCs running Windows XP OS. The PC was equipped with a GPRS modem, a GPS module and a colour touch display with proprietary graphical user interface (GUI) for both the truck and the paver. The connections to the database and the web pages were done over the GPRS.

The PC was connected to the external embedded IO and radio module (Figure 2) through the serial port. The embedded module included a short-range radio and a sensor interface.

2.3. Asphalt mixing plant unit

The asphalt mixing plant unit was a laptop PC with Windows XP. For internet connection a wireless Flash-OFDM technique was used. Flash-OFDM is a proprietary wireless broadband network connection using operation frequency of 450 MHz. It is used only in couple of countries, e.g. Finland, Slovakia and Ireland, because in most European countries the 450 MHz frequency range is preserved for other purposes. The maximum cell radius of the Flash-OFDM is about 30 km. This is significantly longer compared to the 3G mobile phone networks, where cell radius is about 5 km. This makes Flash-OFDM technique applicable in sparsely populated areas. The asphalt mixing plant unit was also connected to the embedded IO and radio module in order to send and receive short range messages. [Arjona, Kerttula & Yla-Jaaski 2008, Tinkler 2006]

2.4 Embedded IO and radio module

The embedded IO and radio module, illustrated in Figure 3 with part numbers, contained a main box (1) and was connected to PCs with the serial bus and build on processor card with ATMEL32 (2) processor, four AD-channels and interface for the NanoNET TRX (Nanotron Tech, GmbH) short-range industrial radio card.

The radio and antenna (3) was earlier tested in VTT projects [Keski-Säntti J., Parkkila T., Leinonen J. & Leinonen P. 2006] and was shown to be usable in difficult environments.

2.5 Sensors

Three different sensing elements were connected to the AD-channels in order to provide process and status data. There were two different types of temperature sensors connected in order to compare and verify the output values. Functionalities of sensors were platinum resistance (PT-100, 4) and infrared (IR, 5).

PT sensors were placed to the bottom of the truck platforms to measure the heat conduction through the platform core and insulated with plywood boards and insulation wool against the cold weather and wind. PT sensors are robust and cheap, but with slow response.

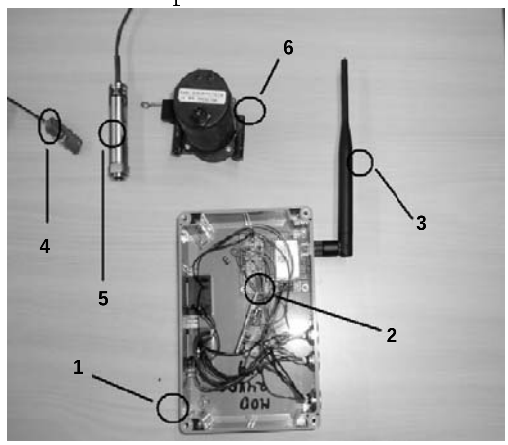


Figure 3. Embedded sensing and radio unit.

Truck IR sensors were placed to a backboard, as shown in Figure 4, and to a special stand in the paver, both pointing to processed mass. IR sensors have fast response, but they are error prone and quite expensive.

Figure 4. IR sensor in the backboard of a truck.

In addition there was a potentiometer based sensing element (6) to detect the operation status of the truck. The potentiometer was connected between the platform and the skeleton of the truck to tell to the system when the unloading was in process.

2.5 Server unit

The system process data was stored to a MySQL type database tables. The database contained five different tables for vehicle, delivery, target, customer and order data. The data was used in order to automate process and offer important information for the operator use.

Other part of the server unit was web pages. There were two different web pages. Other was a complete view of the process units used by asphalt mixing plant unit in a map view with the table providing relevant data of each unit. The process view can be seen in Figure 5.

The other page was also a map and it showed a route and a distance of the corresponding truck to the paver it was supposed to meet. Both maps were read using the Google Maps API. Languages used in web site building were PHP and JavaScript.

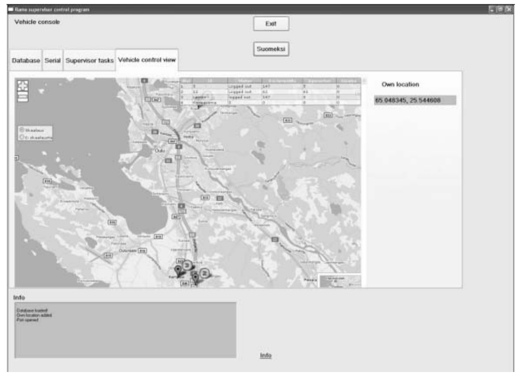


Figure 5. The asphalt mixing plant unit map view.

3. Prototype System Functionality

The functionality of the prototype system was designed to run in parallel with the normal recent system in paving process. The test system contained following operators:

- asphalt mixing station controller
- transporting truck driver
- paver driver

3.1 Features

Identification of trucks using short-range radio

Before all the interactions, receiving vehicle confirmed the identification of the sending vehicle. All the short-range messages contained a field for a vehicle licence number. Using that number, the corresponding vehicle was searched from the database and its location was compared to receiver's own location. If the location was too far away or the corresponding vehicle is not found, the message was then omitted. This RFID functionality was realised using the industrial short-range radios.

Material tracking

Since one of the most important objectives of this project was quality assurance, the material tracking was also improved. During the transportation, the location and the temperature of the mass were saved to a hard drive of the truck PC and the collection of the status data, so called delivery note was saved finally to a database after the unloading. Also the paver was all the time storing the temperature and the location of the mass to the memory. These measures assure that a mass temperature could be checked afterwards in any location of the process from the mixing station to the pavement.

Since the material data was stored all the time to the database with related information, the usage of receipts and paperwork could be reduced.

Functionality automation

Few manual process tasks were also automated in order to streamline the process. Such tasks were for example an order data delivery to the truck driver and a load data delivery to the paver driver.

Virtual delivery note and automatic billing

A concept of automatic bill generation was tested. The bill was assembled using a virtual delivery note containing mass amount, temperature and location data collected during the transportation. The bill was later sent to the customer using the e-mail right after the unloading.

3.2 Communication during work circulation

Figure 6 shows messages exchange between the process participants.

Arriving to the mixing station

When the truck arrives to the mixing station area, the arrival is detected using the GPS module. The short-range radio is then switched on and used for an identification and communication with the mixing station.

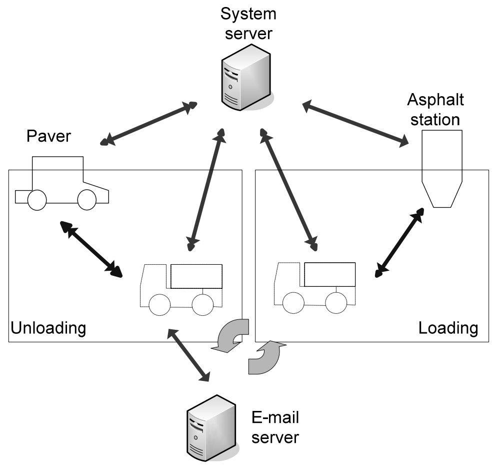


Figure 6. Process message exchange during a work circulation

Loading and weighing

After the truck identification the corresponding order number was sent by the asphalt station controller to the truck. The truck used the order number to get the order data from the database. Based on the order, data was printed to the truck drivers GUI.

Again, based on the order data the truck driver accepted the task and filled in the truck and weighted it. The weighting system sent the data to the database and the required mass variable was updated.

Transporting mass to the paver

After the weighting the truck then started its journey to the correct paver. The mass temperature and location data were logged in 10 seconds interval to a specific delivery note with the order information to store process data for quality assertion. The route view was also updated continuously in same interval in order to guide the driver. Also the truck data to the database was updated simultaneously.

Arriving to the paver

When the truck arrived to the correct paver, similar handshaking like with the mixing station took place. The short-range radio was used and the location of the truck was compared from the database. The truck program also sent the mass status data to the paver and it was printed to the paver driver GUI screen. The paver driver could then be sure of the mass correctness and accept it.

Unloading

If the paver driver accepted the load, the truck driver got the note to the GUI screen and unloaded the mass to the paver. Since unloading was done, the paver driver sent a short-range radio confirmation message to the truck. After the message, the delivery note was stored completely to the database and the summary of the delivery was sent to the customer's e-mail address. In addition, a temperature of paved mass is measured and logged with the location within the paver too.

3.3. User interfaces

Since all the used system platforms were based on PC and Windows XP OS techniques, all the user interfaces were graphic. GUI programs provided information of the process status and enabled interaction between process entities etc.

There were three different GUIs for trucks, paver and asphalt mixer station. Truck and paver versions were designed to be used with touch screen.

Truck GUI

The truck GUI for example showed information of a suggested and performed orders and show information of the transported mass and platform pose that is seen in Figure 7.

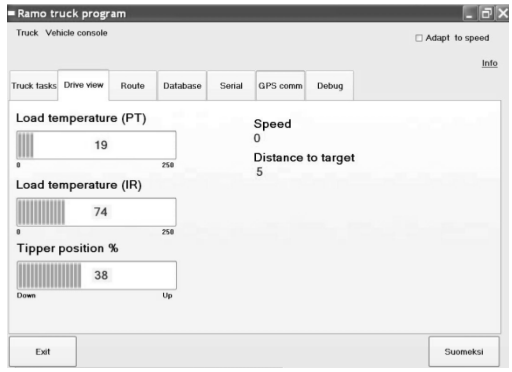


Figure 7. The driver view of the truck GUI

The truck GUI also had a feature of updating route map realised using a help of the Google Maps API. The route map was a web page in the server that received two coordinates and showed the route with the distance between the coordinates and the estimated driving time to destination. The route map is illustrated in Figure 8.

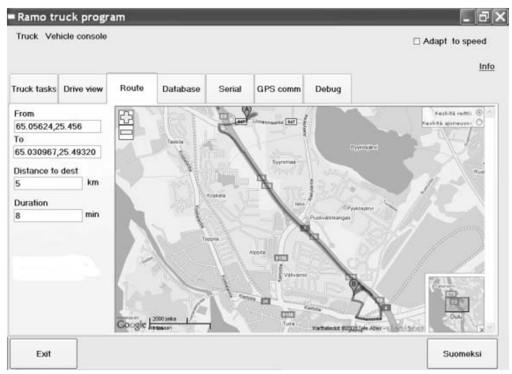


Figure 8. A route in the truck GUI

There was also an adaptive feature in the truck GUI. If the adaptation option was selected, only the most important data was printed to the GUI when the truck speed exceeded 30 km/h.

Paver GUI

In the paver GUI, the map view was also a web page using the Google Maps API. The map showed all the vehicles of the test system with a table of status data like estimated arrival time of the trucks. The paver GUI showed also the mass temperature from the laid mass and the transported mass data from the truck in the case of the unloading.

Asphalt mixing station GUI

The asphalt mixing station's GUI show simply the process view map with all the vehicles.

4. PROTOTYPE SYSTEM FIELD TESTING

A purpose of the system field testing phase was to implement and test the system with a real life environment and equipment. The field tests consisted one paver, three trucks and a mixing station, all instrumented with equipment described in Table 1.

During the tests the vehicles drove and performed their normal process tasks normally and the tests were done in parallel with the normal process flow causing some problems in debugging and data collection.

Since the prototype system was tested in a real environment, there were some malfunctioning and disadvantages found.

4.1 Testing of the wireless communication techniques

The suitability of the communication methods like GPRS, Flash-OFDM and NanoNET was followed.

Some specific performance tests were done for the Flash-OFDM technique.

Table 1. The field test equipment.

	Sensors	**Radios**
Asphalti station	-	Flash-OFDM, NanoNET
Truck 1	IR, PT, tipper	GPRS, NanoNET
Truck 2	PT, tipper	GPRS, NanoNET
Truck 3	PT, tipper	GPRS, NanoNET
Paver	IR	GPRS, NanoNET

GPRS

Also the GPRS based internet connection caused some problems during the tests. The weak signal in rural areas was not fast enough for some functionalities and connection was lost every now and then too.

Flash-OFDM

Performance of the Flash-OFDM technique was tested on a moving vehicle. Average ping time to the VTT's server was 69 ms, typical download speed was between 500 – 800 kb/s and upload speed was between 180 – 250 kb/s. The speed of the vehicle was about 100 km/h. On areas were Flash-OFDM network coverage is good this technique offers reliable internet connection for moving vehicles.

NanoNET

A vehicle to vehicle communication was done using the robust NanoNET TRX short-range industrial radio. The range was between 20 and 200 meters depending on a used antenna type.
Different kind of antennas was tested in order to improve the message delivery between the NanoNET terminals. The antennas that were originally used were too directional and the narrow close-field reached too far. These features caused problems in close-range communication and RFID identification. Problems could be solved with omni-directional antennas.

4.2 Asphalt mass temperature measurements

One very interesting topic of the tests was a correctness of the sensor values. Both IR and PT sensor values were logged and compared together. Figure 9 shows the mass temperature comparison of truck load between the different sensors during a load transport. As it can be seen the PT sensor with the black curve reacts slower to the mass temperature changes. Moments of unloading can be seen from the green tipper pose curve where non-zero values show the unloading timings. The unloading of a one load is usually done in several phases, also in the case.

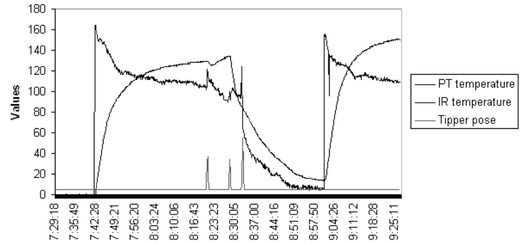


Figure 9. The functionality of temperature sensors.

Mass temperatures from the paver IR sensor was measured and compared to values from a handheld Raytek MX4 thermometer. Figure 10 shows the temperature values in a time window. Handheld thermometer values point also out the times when the unloading of the certain truck load starts.

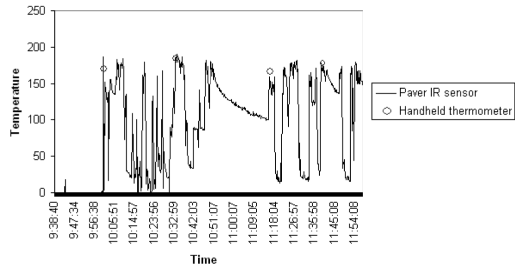


Figure 10. The paver IR sensor and thermometer temperatures.

For the further research and PT sensor ideal location, thermographic camera pictures were also taken. More exact heat distribution of the load could be seen from these pictures. Figure 11 shows an example of these pictures.

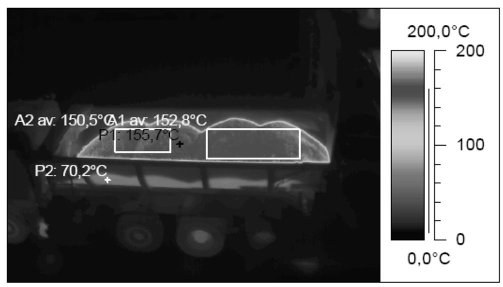


Figure 11. A thermographic camera image after a truck loading.

4.3 Usability of the system

There were plenty of specific topics that were evaluated during the tests. The stability of the GUI programs and actions performed within were naturally vital. Also the usability of the few needed operator interventions in the GUI program during the process flow were tested by asking about use experiences from the drivers.

Drivers are clearly willing to adopt new technology, but it could be seen that this kind of system must be designed to be used with as few user interventions as possible.

5. Conclusions

In this paper, the system, that can be used to control asphalt paving process, was demonstrated. This control process requires for real-time follow-up that involves data collection, storing and using in order to automate and control the paving process to gain new benefits.

Wireless communication is an enabler for collecting data in real time from a construction process. VTT has studied different approaches for this. Mobile phone technology was used in "Wireless construction site" – project, where a system for managing mass transportation tasks in road construction was developed and tested. [Rannanjärvi L. 2003] The main advantages of using mobile phone technology are low cost portable devices. This makes easy to take the system into the use. However, in this approach it totally depends on the user how reliable and up-to-date the collected data is.

Another approach was to use embedded platform with sensors to enable process management tasks. [Kilpeläinen, Heikkilä & Parkkila 2007] Embedded systems are cheaper and easier to install compared to complete in-vehicle PCs, but on the contrary they don't provide that much possibilities concerning especially user interfaces.

This paper presents more automated data collection methods using sensors installed to the work machines with PCs. The most obvious disadvantage is that constant installations to the machines are needed. On the contrary some of the data collection tasks (e.g. material tracking) can be automated so that minimum effort is needed from the user. In practise a combination of interactive user based and automatic sensor based data collection is the most practical approach.

The large graphical touch-screens in vehicles fulfilled its task very well providing essential information of the process. Also the use of RFID –like identification helped to automate some tasks and provided also optional communications link between vehicles.

One big issue is what to do with all the data collected. Of course the data can be stored to the database and used for quality assurance and reporting purposes. In this project, that data was also used in real-time to control the process. The critical process data is displayed to the process operators and some of the decision-making process is transferred from the operators to the programs.

The prototype system designed and tested in this project in fall 2008 supports a conclusion that there could be plenty of process improving acts done in the asphalt paving process. The quality control could be made much more efficient and accurate. The vehicle guidance could be improved much too with the security matters. Also the process safety could be improved with help of technical solutions. One of the most important perception is that the prototype system improves the functionality not only the certain construction site but the all the sites that gets the asphalt mass from the certain asphalt mixer.

Since PCs in professional vehicles are getting more and more common, the system like the tested one could be implemented easier and cheaper that makes it extremely reasonable to develop it further. Wireless technologies are developing rapidly and broadband internet access can be available to the vehicles and work machines more commonly. Also existing vehicle bus architectures like CAN should be exploited more vastly when connecting external sensors and terminals to the system. Truck load scale could also enable complete automation to the billing since the sub-loads could be measured in a case where one truck-load would be dispensed to different locations.

References

[1] Arjona, A., Kerttula, J. & Yla-Jaaski, A. 2008. Live Network Performance Challenge: FLASH-OFDM vs. HSDPA. AINA 2008
The IEEE 22nd International Conference on
Advanced Information Networking and Applications. Naha City Okinawa 900-0015 JAPAN, 25-28 March 2008. Pp. 918.

[2] Keski-Säntti J., Parkkila T., Leinonen J. & Leinonen P. 2006. Wireless communication and mems sensors for cheaper condition monitoring and prognostics of charging crane. Proc. of the 19th International Congress of Condition Monitoring and Diagnostic Engineering Management, COMADEM 2006. Lulea, SWE, Pp. 747.

[3] Kilpeläinen , P., Heikkilä , R. & Parkkila , T. 2007. Automation and wireless communication technologies in road rehabilitation. Proceedings of the 24th International Symposium on Automation and Robotics in Construction (ISARC). 01/01/2007. Pp. 35.

[4] Lu, M., Chen, W., Shen, X., Hoi-Ching Lam & Liu, J. 2007. Positioning and tracking construction vehicles in highly dense urban areas and building construction sites. 647-56 p. ISBN 0926-5805.

[5] Nassar, K., Thabet, W. & Beliveau, Y. 2003. Simulation of asphalt paving operations under lane closure conditions. Automation in Construction, 9, Vol. 12, No. 5, pp. 527-541. ISSN 0926-5805. doi: DOI: 10.1016/S0926-5805(03)00005-0.

[6] Peyret, F., Jurasz, J., Carrel, A., Zekri, E. & Gorham, B. 2000. The Computer Integrated Road Construction project. Automation in Construction, 9, Vol. 9, No. 5-6, pp. 447-461. ISSN 0926-5805. doi: DOI: 10.1016/S0926-5805(00)00057-1.

[7] Rannanjärvi L., K.K. 2003. Remote control and monitoring over Internet - Wireless Construction. 7th IFAC Workshop on Intelligent Manufacturing Systems. Budapest, HU, 6. - 8. Apr. 2003. Pp. 205.

[8] Tinkler, R. 2006, 13/02/2006-last update, QUALCOMM announces shipment of second-generation FLASH-OFDM platform for mobile broadband access [Homepage of Qualcomm], [Online]. Available: http://www.qualcomm.co.in/press/releases/2006/060213_shipment_second_generation.html [2008, 12/03] .

Application of Appropriate Technologies to 3D Local Terrain Modeling in Real-time for Intelligent Excavating System (IES)

Byung-In Yu[1], Hyun-seok Yoo[2] , Young-suk Kim[3], Jong-won Seo[4] and Seungwoo Han[5]

[1]Graduate Student, Department of architectural Engineering, Inha University, Incheon, 402-751, South Korea; PH (82) 32-872-9757; FAX (82) 32-866-4624; e-mail: ybi2222@gmail.com
[2]Ph.D. Candidate, Department of architectural Engineering, Inha University, Incheon, 402-751, South Korea; PH (82) 32-872-9757; FAX (82) 32-866-4624; e-mail: hsyoo.inha@gmail.com
[3]Associate Professor, Department of architectural Engineering, Inha University, Incheon, 402-751, South Korea; PH (82) 32-860-7593; FAX (82) 32-866-4624; e-mail: youngsuk@hotmail.com
[4]Associate Professor, Department of Civil Engineering, Hanyang University, Seoul, 133-791, South Korea; PH (82) 2-2220-1482; FAX (82) 2-2220-4480; e-mail: jseo@hanyang.ac.kr[5]Assistant Professor, Department of architectural Engineering, Inha University, Incheon, 402-751, South Korea; PH (82) 32-860-7595; FAX (82) 32-866-4624; e-mail: shan@inha.ac.kr

Abstract

Automation in construction is considered as one of solutions to fundamental problems which the construction industry has confronted. The development of real-time 3D local terrain modeling is the core methodology of achievement of IES (Intelligent Excavating System) which is a mega-size research concerning automation in construction. Stereo vision technique was selected as an appropriate method to achieve 3D terrain modeling in real-time. The objective of this study is to suggest a conceptual method by use of the stereo vision and assess feasibility of practical application under various conditions of construction sites. Bumblebee XB-3 is adopted as the stereo vision system in a prototype of a 3D terrain modeling system. This study is to investigate fundamental theories of the stereo vision system and to conduct tests in indoor and outdoor environments for determining appropriate detection range, finding installation location of cameras, doing visual assessment of the technical appropriateness, and checking reliability of 3D visual data. These decomposed tasks present valuable information for establishment of the conceptual method and improvement on the limitations of the use of stereo vision system for 3D terrain modeling in real-time.

Keywords: automation, stereo vision, 3D, terrain model

1. Introduction

Many construction planners and site engineers have regarded enhancement of construction performances by high construction productivity and low unit costs as one of crucial factors for determining a success of construction projects. The shortage of experienced laborers, raise of labor costs, and debasement of construction qualities make construction industry in Korea face serious difficulties in achievement of high construction performances. Construction industry has tried to use emerging technologies that are capable of providing excellent construction performances. These efforts are especially witnessed in a field of construction equipment. A research project titled "Intelligent Excavating System (IES)" that is one of research projects supported by Korean government started with professionals in industries and researchers in academic institutes in Korea at 2007 is currently conducting (Ahn et al. 2008; Yu et al. 2008).

Stereo vision technique was selected as an appropriate method to achieve 3D terrain modeling in real-time in previous researches (Ahn et al. 2008; Yu et al. 2008). The objective of this study is to suggest a conceptual method by use of the stereo vision and assess feasibility of practical application under various conditions of construction sites. In this study, Bumblebee XB-3, a state-of-the-art stereo vision system, is adopted as the stereo vision system which is used in a prototype of a 3D terrain modeling system. In accordance with the goal of the research, this study is to investigate fundamental theories of the stereo vision system and findings by past researches related to this study. This study is to conduct tests in indoor and outdoor environments based on theoretical reviews. Indoor test allows us to determine an appropriate detection range by the stereo vision system and gives us basic information about installation location and installation angle of cameras on a cabin of an excavator. It would be primary information for establishment of a conceptual method for practical application. Outdoor test enables to do visual assessment of the technical appropriateness of the stereo vision system for practical application to construction sites. Accuracy test, another outdoor test, provides quantified differences of distance data of 3D images by the stereo vision system with comparison of the result by actual measurement. These decomposed tasks mentioned previously present valuable information for establishment of the conceptual method and improvement on the limitations of the use of stereo vision system for 3D terrain modeling in real-time.

2. Literature Reviews

Stereo vision technique is currently used in a broad range of various fields such as terrain model acquisition, object identification and tracking systems. This technique is implemented by two identical cameras placing separately, that is, those are placed on different horizontal level but the same vertical level (Yu et al. 2008). The pictures taken by two cameras are merged and produce a virtually instantaneous estimate of the distances to objects. These cameras detect images of an identical object with different horizontal locations and different distances. These two detected images are merged based on an identical matching point of the object, a corresponding pixel, and then creates one merged image. This merged image presents distance data from cameras to the object. Quick measurement in distance yields that this technique can be highly effective for detecting the object and segmenting the object and gauging its size and shape (Woodfill et al. 2007; Cho and Park 1997; Ahn et al. 2008; Yu et al. 2008).

Identifying distance data from the camera to the object is derived from figure 1 and the equation as below.

$$z = f \times \frac{b}{d} \qquad \text{: Equation (1)}$$

where, z: distance from the camera to the object point

f: focal length

b: baseline

d: disparity between the distances from the left image and the right image

$d = d_l - d_r$

d_l: the distance from the left image to the object point

d_r: the distance from the left image to the object point

As described the equation (1), computation of disparity based on the corresponding pixel in two images, the left image and the right image would be a technical issue in stereo vision technique.

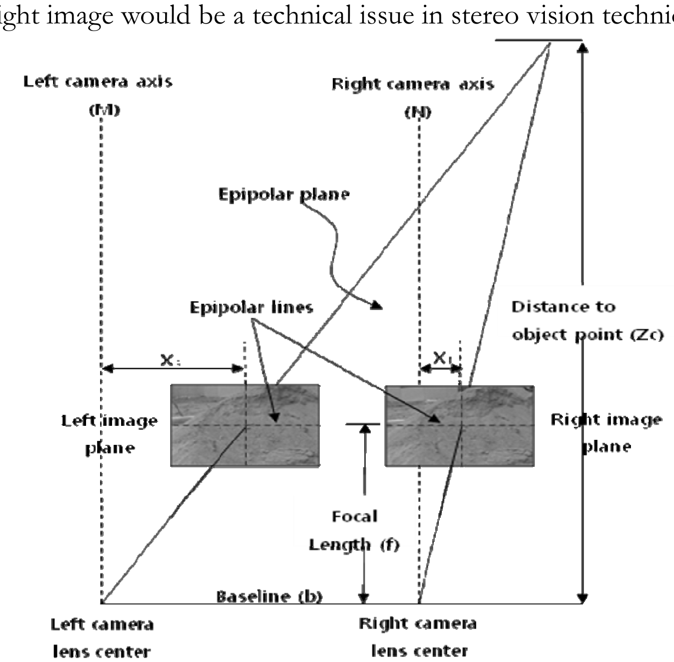


Figure 1. Diagram of basic operation by stereo vision technique

In addition to computation of the disparity, it would also be critical issue to find an appropriate corresponding pixel in an identical object with two images. Two methods have been introduced on the market: 1) feature-based method, and 2) area-based method to resolve this difficulty.

Primary conditions for achieving the objective of this study is to detect a terrain image in real-time without any interruption of equipment operation. One cycle of excavating is composed of "scooping out" and "unloading the earth". It generally takes 10 seconds for completing one cycle according to several equipment guide books and site observation. This information notifies us that all procedure including detecting and image processing would be completed within 10 seconds. Detecting terrain image is another limited condition in this study. Detection by cameras would finish as fast as it can so that it would not give any interruption during scooping and unloading the earth on the tray of trucks. These two limitations described previously would be resolved by technical specifications of state-of-the-art stereo vision system. The emerging technique of stereo vision system provides the technical solutions in that interpreting visual data to 3D from distance data takes generally 1 ~ 2 seconds and detecting the image takes milliseconds.

3. Conceptual Method of Stereo Vision for 3D Terrain Modeling

Appropriate conceptual method of stereo vision system requires that detection range of cameras and horizontal excavation range would be primarily determined. Careful study on installation of stereo vision cameras follows. These requirements make this study investigate the detection range of cameras by indoor tests and the excavation range by construction equipment references. The investigation allows this study to determine the appropriate installation location of the stereo vision cameras.

3.1. Estimation of detection angle

Available detection range of the stereo vision camera is delivered by an indoor test. According to the technical specification of Bumblebee XB-3, available angles are vertical viewing angle of 49.5 degree and horizontal viewing angle of 66 degree. Indoor test is required so that appropriate viewing angle for use in excavation operation. The system created 3D images based on two images captured by cameras which were located in 2 meter distance from a screen. The size of projected area in screen was 2.28 meter in width and 1.75 meter height. This result and proportional computation informed that appropriate angles by the stereo vision system were vertical viewing angle of 47 degree and horizontal viewing angle of 59 degree.

3.2. Installation of the stereo vision system

An excavator which is under development as a prototype shows 8.2 meter of maximum reach and 5.0 meter of maximum depth below ground. Previous study (Seo et al. 2008) indicated that appropriate horizontal excavating range was 3.4 ~ 7.67 meters when the specification data of the same excavator were delivered.

Location and mounted angle of the stereo vision system are determined by information derived from appropriate detection angle and horizontal excavating range delivered previously. On-site observation of excavating operation informed that the stereo vision system was installed on a cabin of the excavator to obtain effective detection with appropriate coverage of the excavated area. Specification of the excavator used in this study says that that height of prospective installation location on the cabin is 2.8 meter. Figure 1 illustrates basic simulation chart of the excavator for searching appropriate installation location of the stereo vision system.

Figure 2 indicates that the stereo vision system needs to be installed on the cabin with 2.8 meter height and 30 degree incline to cover 3.4 ~ 7.67 meters of appropriate horizontal excavating area.

4. Application Analysis by Field Tests

4.1. 3D terrain modeling

There are a few of specific technologies such as Triclops SDK by Point Grey Research Inc., Open GL (Open Graphic Library) by Silicon Graphics Inc., and Microsoft Visual C^{++} compiler, version 6.0 by Microsoft Corp. which are utilized for 3D terrain modeling. Triclops SDK allows us to control the operation of stereo vision cameras and to transmit image data. It also includes conversion function from the image data to point cloud suing SAD (Sum of Absolute Difference) algorithm. Open GL makes the converted point cloud by Triclops SDK be illustrated in the user interface. Figure 3 presents basic diagram of functions by Triclops SDK and Open GL.

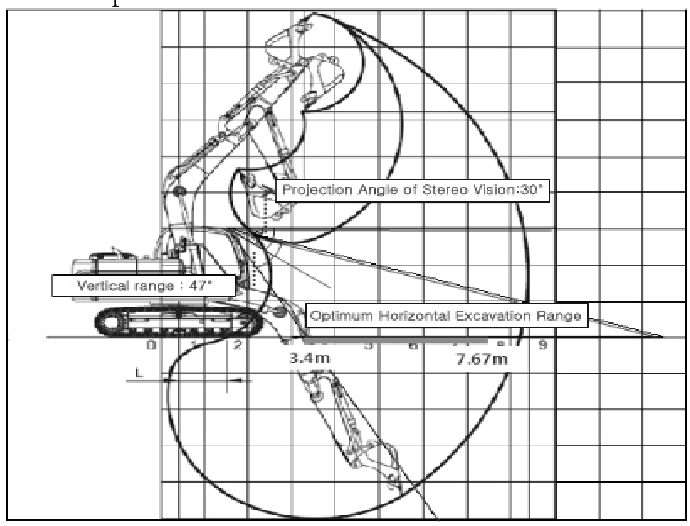


Figure 2. Operation chart of the excavator

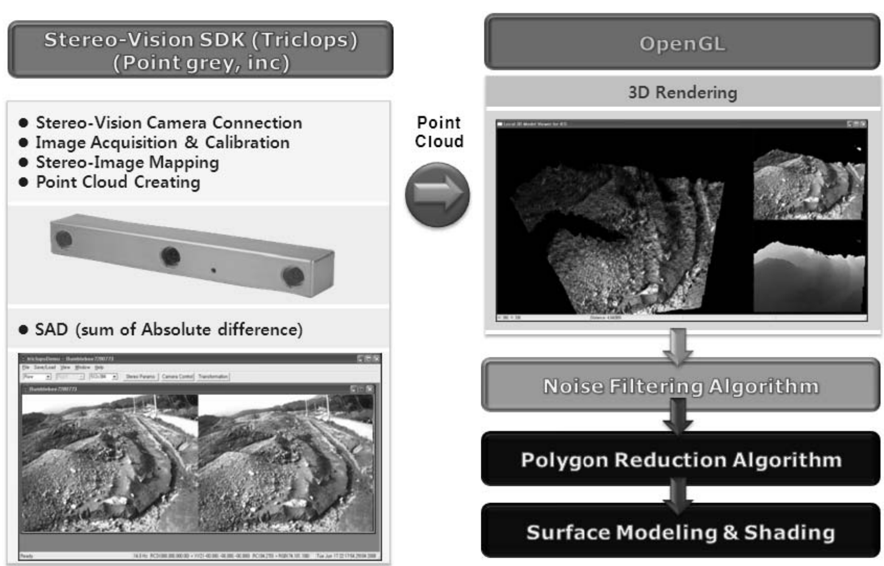


Figure 3. Diagram of functions by Triclops SDK and Open GL

The images detected by stereo vision cameras show 640× 480 resolutions in color and are merged based on the image captured by the camera on the right location. Figure 4 shows a 3D modeling display viewer which is composed with 3D terrain model display, 2D stereo image map created by matching two captured images, and a depth map illustrated as a grey scale index which is converted from the computed distance data. The conceptual method of stereo vision system conducted by this study enables the camera to detect the terrain image on the trigger signal by excavating system automatically or by manually. The merged image after matching then is automatically presented on the 3D modeling display viewer.

4.2. Field tests

Many field tests on various construction sites are conducted to assess the practical application of stereo vision system for 3D terrain modeling. Various shapes of terrain and objects which may exist in construction sites are detected in field tests.

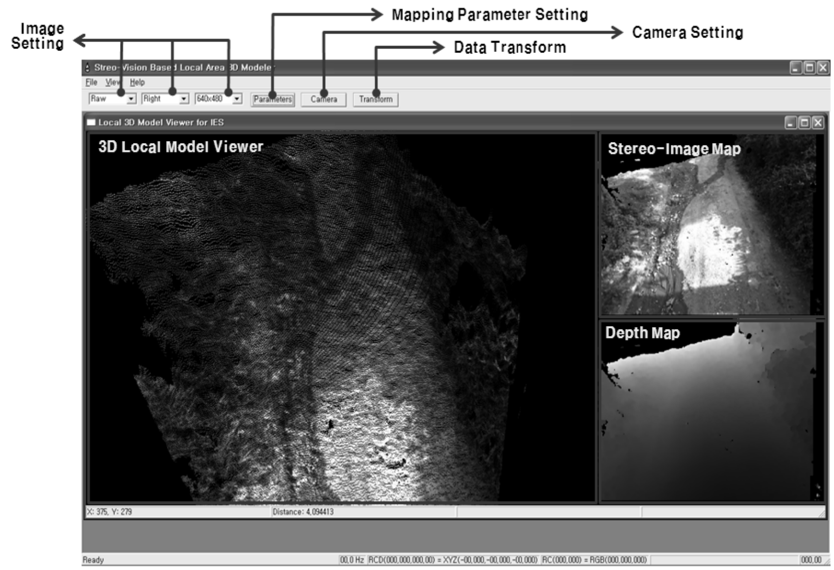


Figure 4. 3D modeling display viewer

The stereo vision system is mounted on a tripod of 2.8 meter height and 30 degree incline from the horizon. The images is detected with 640× 480 resolutions in color. Total 98 images are captured by total 5 field tests. Basically, all captured images are converted to 3D image via a specific display system which was developed in this study. Consequently, assessment of noise conditions and processing time makes the conceptual method of stereo vision system feasible or improved further.

Figures 5 and 6 show 3D terrain images of various shapes of terrain and with various types of objects existing in construction sites, respectively. All 3D images converted from the captured images shown in Figures 4 and 5 are assessed and determined whether these are sufficiently appropriate to be applied in excavating works.

Professionals agreed that most 3D images showed appropriate accuracy with consideration of a broad range of operation tolerance of excavation. This favorable test result is due to easiness of finding corresponding pixels in terrains while detecting various patterns in terrain.

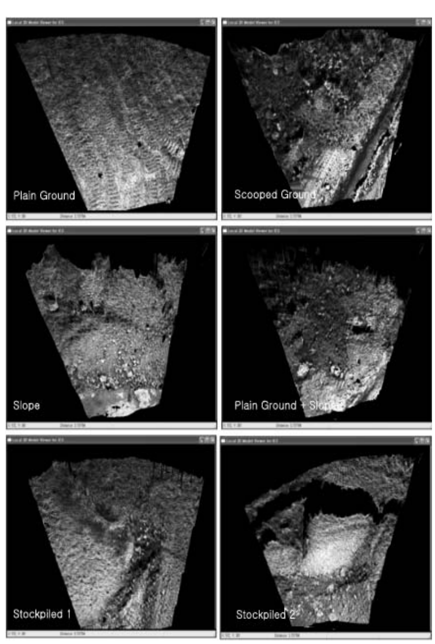


Figure 5. 3D terrain images of various shapes of terrain

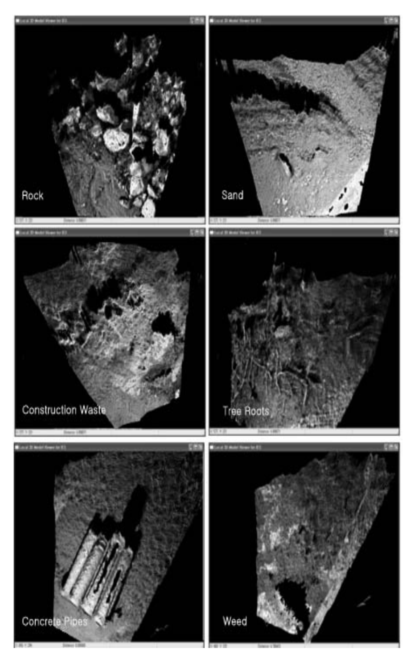


Figure 6. 3D terrain images with various objects

Unlike most 3D images, the images that concreted hume pipes and surface water were captured in are determined not to be used due to poor quality caused by not finding corresponding pixels. This limitation would be resolved by controlling camera's shutter speed and creating noise removal algorithm. Figure 7 shows the 3D images of terrain with concrete hume pipes and surface water in poor quality

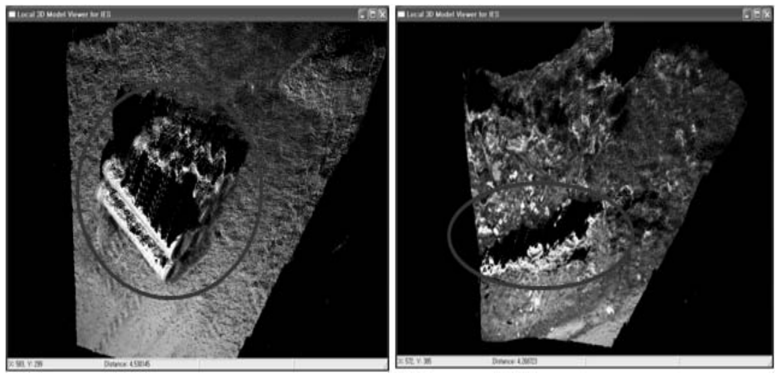


Figure 7. 3D terrain images with concrete hume pipes and surface water

Visual assessment shown in figures 4, 5, and 6 show proper result in accuracy. It doubts whether 3D images created by stereo vision system are perfectly matched with actual terrain information or not. Consequently, the accuracy test requires that the distance data in 3D images by stereo vision system need to be compared with those by actual measurement.

Figure 8 shows an accuracy test conducted in construction sites.

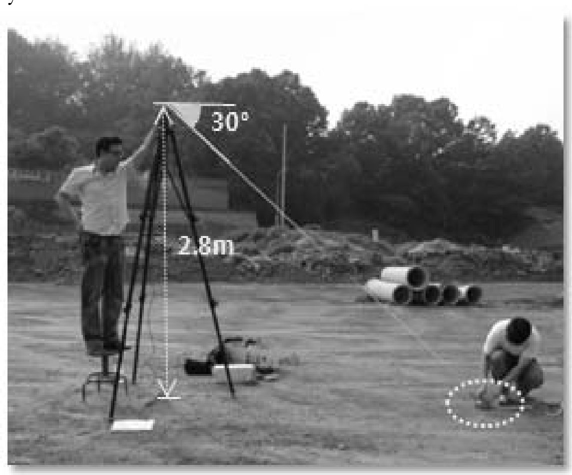


Figure 8. Accuracy test of stereo vision system

As shown in figure 8, the accuracy test is conducted that the stereo vision system which is located in 2.8 meter height and 30 degree incline detects targets which placed randomly within 2 ~ 11 meters. The distance data by stereo vision system are compared with those by actual surveying in that the distance from the camera to the target is measured manually.

Comparison of total 148 samples of the distance data between 3D images and actual measurements indicates 0.051 meter of an average deviation and 0.168 meter of a maximum deviation. Graphs where deviations analyzed in the 4th and the 5th field tests were illustrated in figure 8. In addition to deviation graphs, table 1 presents analysis results of deviation by distance from the camera and the target.

As shown in figure 9 and table 1, it notes that more distance from the camera to the target has bigger deviations of the distance data between the 3D images and actual measurement. It explains that the stereo vision system has limitation for practical application to detecting terrain where is located with long distance from the camera due to difficulty in finding the corresponding pixel.

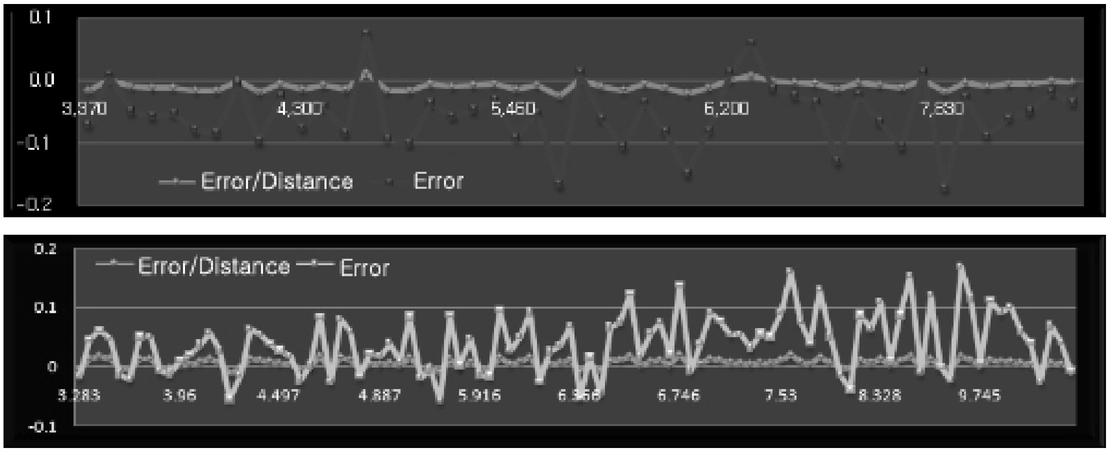


Figure 9. Deviation graphs in the 4th and the 5th field tests

Table 1. Comparison result of deviations (unit: meter)

Distance range	Maximum deviations	Average deviations	Standard deviation	Variance
3 ~5	0.083	0.034	0.0213	0.0005
5 ~7	0.136	0.050	0.0336	0.0011
7 ~9	0.158	0.068	0.0431	0.0019
9 ~11	0.168	0.065	0.0484	0.0023

5. Conclusions and Future Works

This study introduced the conceptual method of stereo vision system to be applied to create 3D terrain model in excavation operation. The stereo vision system has been deployed in broad ranges of area where it is normally limited to stationary objects in outdoor environments or steady-state objects in indoor. The application of stereo vision system to construction operation for creating 3D terrain model would face many challenges which are totally different with past researches. Accordingly, this study conducted many experiments using stereo vision system in construction sites in order to achieve appropriate applications and assessed technical characteristics including limitations. These experimental results suggest conclusions as follows:

- Indoor experiment conducted to find appropriate installation of stereo vision system on cabin of an excavator. Experimental results inform that the camera on the excavator is located in 2.8 meter height and 30 degree incline for acquiring available detecting areas.
- As result of several field tests in construction sites, most targeted terrain is well detected and created in 3D images. A few of images detecting concrete hume pipes and surface water have a problem that white hole is witnessed due to mismatching two images. It, however, would be overcome with use of more advanced algorithm for noise removal. The accuracy test of the distance data of 3D images showed that average deviation of distance data between 3D images and actual measurement was 0.051 meter. These deviations were gradually bigger while the distance from the camera to the target became more.
- In this study, the conducted tests of accuracy of 3D images by stereo vision system were limited only to visual assessment of 3D images and comparison assessment of distance data between 3D images and actual measurement. Although these tests provided basic guideline to the conceptual method, well-designed experiments need to be continuously developed. Accuracy and stability of this system on vibration by an excavator and reliability and accuracy of the acquired 3D images would be further experimented.

The primary findings and conceptual method suggested in this study would be applied to not only an excavator but also a dozer, a loader and other construction equipment. Based on findings in this study, further studies for achieving accurate 3D terrain model would be conducted in IES research project.

Acknowledgement

This research was supported by a grant from the research titled "Development of Intelligent Excavating System", a part of Construction Technology Innovation Program(CTIP) by Ministry of Land, Transportation and Maritime Affairs(MLTM) of Korean government.

References

[1] Ahn, J. S., Yu, B. I., Kim, Y. S., and Han, S. W. (2008). "Emerging Technologies for Real-time 3D Terrain Modeling and Automatic Soil Volume Computation in Earthmoving Operation" Proc., 12th ICCCBE &2008 ICITC, 276.
[2] Cho I. S., and Park J. G. (1997). "Extraction of Depth Information using Correspondence Analysis in Stereo Vision." Basic Science and Engineering, 1(2), 899-905.
[3] Lu, M., Chen, W., shen, X., Lam, H., and Liu, J. (2007) "Positioning and tracking construction vehicles in highly dense urban area and building construction sites." Automation in Construction, 16(5), 647-656
[4] Woodfill, J. I., Buck, R., Jurasek, D., Gordon, G., Brown, T. (2007). "3D Vision: Developing an Embeded Stereo Vision System." IEEE Computer.
[5] Yu, B. I., Ahn, J. S., Lee, J. H., Yu, H. S., Kim, Y. S., and Han, S. W. (2008). "Implementation of Real-time 3D Terrain Modeling in Automatic Earthmoving Operations." Proc., 11th Symposium on Construction Robotics in Japan, 75-80.

Optimum Control of Vibratory Piling Process

Michel Cotsaftis[1] and Erno Keskinen[2]

[1]ECE, 37 Quai de Grenelle,75725, Paris Cedex 15, France, Email : mcot@ece.fr
[2]Laboratory of Machine Dynamics, TUT, P.O.Box 589, FIN-33101 Tampere, Finland, Email: erno.keskinen@tut.fi

Abstract

A difficult task in densely constructed urban areas is to proceed to new extensions or to unavoidable repairs without creating damages to existing buildings. The method so far is to isolate the worked zone by sheet piles to maintain constant earth pressure conditions against existing foundations. But even if they are much less harmful than classical hammering methods, the vibratory methods in use to drive the sheet piles into the ground are still producing oscillations which propagate into the ground and may endanger the neighbouring buildings especially if a resonance phenomenon do occur. A control method is analysed in the following to keep the vibration amplitude below a fixed safety level while the sheet itself is piled in minimum time into its pre-assigned position.

Keywords: Vibratory Piling, Sheet Piling, Ground Vibrations

1. Introduction

Urban constructions are typically densely packed to accommodate high population density, and are very often located over relatively soft soils such alluvial ones, reminiscent of the early town origin close to water source. A consequence is that building renewing and/or repairing now requires to lay down very deep foundations to cast pillars in or to hit directly piles into the ground. This operation in turn creates serious difficulties as, even with vibratory units which are much less harmful for the environment than classical hammering methods, there is still generation of ground vibrations which propagates to neighbouring buildings and may seriously damage them especially if resonances are produced during the process, corresponding to critical speeds of the global pile-ground-building system [1]. Vibratory units in use are consisting of pair of shafts driven by electrical or fluid actuators to comply with the needed high power density, on which eccentric masses are mounted to excite a vertical harmonic force pushing the sheet or the pile into the ground. With a static driving force on top, the sheet penetrates with a speed depending on vibration amplitude and frequency, and the efficiency is increased by adding several shafts in parallel. With already two pairs, it is then possible to modulate the composition of individual shaft pair forces by adjusting both the phase angle between primary and secondary shafts and the rotation speed of the primary shaft. However, the interaction dynamics between the various sub-systems are becoming more complex, and in an industrial set-up, an automatic monitoring system has to be developed to assist the operator who cannot be left with full driving operation. Useful conditions are that the oscillation amplitudes produced by vibratory piling at fixed neighbouring locations are always below some pre-assigned threshold safety values, and also that the piling time is minimized for economic reasons. Because they are antagonistic, the previous conditions require first to set up the complete system dynamic equations. The corresponding optimum problem is analysed afterward, the properties of required controller are discussed and its analytic expression is given in specific case.

2. System Equations

For a two-shaft line system shown on fig.1, the degrees of freedom are the vertical position z of the vibratory unit of mass m, the position Z of the upper frame of mass M, and the angular positions φ_j of mass eccentricity m_j and inertia moments I_j ($j=1,2$). The dynamical equations then read in normalized form [2]

$$\ddot{Z} + \omega_M^2 (Z - z) = -\omega_g^2 - f_a$$
$$\ddot{z} + \omega_m^2 (z - Z) = -\omega_g^2 - f_s \,\mathrm{sgn}(\dot{z}) + f_p + \sum_{j=1,2} \mu_j [\dot{\varphi}_j^2 \sin(\varphi_j) - \ddot{\varphi}_j \cos(\varphi_j)]$$
$$\ddot{\varphi}_j \pm \omega_j^2 F(\varphi_1 - \varphi_2) = t_j - \mu_j (\omega_g^2 + \ddot{z}) \cos(\varphi_j) \qquad (1)$$

where here $z \to z/l$, $Z \to Z/l$, $\omega_M{}^2 = k/M$, $\omega_m{}^2 = k/m$, $\omega_j{}^2 = K/I_j$, $\omega_g{}^2 = g/l$, $\mu_j = (r/l).(m_j/m)$, $f_a = F_a/Ml$ $f_s = F_s/ml$, $f_p = F_p/ml$ are respectively the applied feeding force, the skin f riction and the point load of the piled object divided by the corresponding mass and the length l of piled object. $F(x)$ is the dead zone function

$$F(x) = x + \pi/2 \qquad \text{if} \qquad x < -\pi/2$$
$$F(x) = 0 \qquad \text{if} \qquad -\pi/2 \le x \le \pi/2 \qquad (2)$$
$$F(x) = x - \pi/2 \qquad \text{if} \qquad x > \pi/2$$

Furthermore the angular acceleration t_j produced by the hydraulic actuators are given by $t_j = \theta\,(p_j^+ - p_j^-)$, with

$$\dot{p}_j^+ = u_j \sqrt{p_s - p_j^+} - \lambda\dot{\varphi}_j$$
$$\dot{p}_j^- = u_j \sqrt{p_j^- - p_t} + \lambda\dot{\varphi}_j \qquad (3)$$

p_s , p_t the normalized source and tank pressures and θ, λ two normalizing parameters. Previous numerical analysis of eqns(1,2,3) for realistic industrial equipment [3] has shown the adequacy of both the model and the very fashionable used software to represent system dynamics. In a next step, the present complement paper is intended to study analytically global guideline properties for optimal control of the system in order to ease its numerical resolution. Along this view, it is first interesting to note that, though the physical control parameters are the u_j which are ultimately governing the inputs to the proportional control flow valves, the total system (1,2,3) could be split into the piled object dynamics from eqns(1,2) with inputs t_j , and the realization of the required inputs from fluid power eqns(3). The only difficulty is in the fact that the intermediate controls t_j satisfying the conditions may not be accessible from the set of possible physical controls u_j, in which case only an approximate optimum will be reached. Then the problem would amount to determine the inputs f_a and t_j so that the piled object reaches its final position inside the ground in the vibratory shortest time while the produced vibration amplitudes at assigned places are during this time always bounded above by a fixed threshold value.

3. Discussion

As just mentioned, the dynamic system (1,2,3) could naturally be split into the dynamic part of eqns(1,2) and the fluid power part of eqns(3). However another possible splitting is obtained by observing from the third equation in system (1) that the first two equations in system (1) are coupled to the set of the third of eqns(1) plus eqns(3) through the only term $\mu_j \ddot{z} \cos(\varphi_j)$, which can be a "small" one . So it is advisable to use this system property and to split instead the complete system in the following way

$$\ddot{Z} + \omega_M^2 (Z - z) = -\omega_g^2 - f_a$$
$$\ddot{z} + \omega_m^2 (z - Z) = -\omega_g^2 - f_s \,\mathrm{sgn}(\dot{z}) + f_p + U \qquad (5)$$

and

$$\ddot{\varphi}_j \pm \omega_j^2 F(\varphi_1 - \varphi_2) = \theta(p_j^+ - p_j^-) - \mu_j \omega_g^2 \cos(\varphi_j) + C_j \qquad (6)$$

$$\dot{p}_j^+ = u_j \sqrt{p_s - p_j^+} - \lambda\dot{\varphi}_j$$
$$\dot{p}_j^- = u_j \sqrt{p_j^- - p_t} + \lambda\dot{\varphi}_j \qquad (7)$$

where

$$C_j = \mu_j \ddot{z} \cos(\varphi_j) \qquad U = \sum_{j=1,2} \mu_j [\dot{\varphi}_j^2 \sin(\varphi_j) - \ddot{\varphi}_j \cos(\varphi_j)] \qquad (8)$$

are respectively the coupling term between the two sets (5) and (6,7), and the input to the first set of eqns(1). This display clearly shows the important role of the rotation frequency of the masses in vibratory units, both with their actual values and with their variations. So in eqns(1) for a fixed applied force f_a the local input U plays the role of a control force which modulates the ground penetration of the piled object, especially when observing from experiments that the reaction force f_p from the ground is generally depending on vibratory unit rotation frequency and typically decays monotonically above some threshold value Ω_c .

With these elements, a very simple optimum bang-bang type running operation would consist in turning the vibratory unit to its maximum rotation speed Ω_{max} to reduce f_p while the frequency shift is driven to the largest vibratory amplitude by fixing the φ_j at the maximum of the source term U, then insuring the fastest penetration in the ground with smallest optimal vibration amplitude at stationary regime. From eqn(8), and writing $\varphi_1 = \varphi$, $\varphi_2 = \varphi + \Delta\varphi$ for simplicity the maximum of U is obtained for $\Delta\varphi = 0$. This procedure would work as long as, during the piling operation, the resulting ground oscillation amplitude at specific sensitive location near by the piling place and propagated from it, is below a pre-determined threshold fixed by a risk of damage. This is not always guaranteed because one should ride over all frequencies in the interval $[0, \Omega_{max}]$ when departing the piling work from rest and one may cross some resonance or simply some sensitive frequency for which the ground response amplitude overpasses the threshold value. The problem is then to downgrade the previous optimum when approaching these frequencies in order to account for the imposed threshold constraint. Noting that $U = 0$ for $\Delta\varphi = \pi$, and that U is monotonic in the interval $[0,\pi]$, this is always possible by manipulating the phase shift between vibratory units. So a natural control strategy is to drive the phase shift so that, in accordance with predetermined amplitude condition all along the frequencies in the operation interval $[0, \Omega_{max}]$

$$A_G(\omega,\Delta\varphi) \le A_{max}(\omega) \quad \text{for} \quad \omega \in [0, \Omega_{max}] \qquad (9)$$

where $A_G(.,.)$ is the resulting ground amplitude at prescribed locations and $A_{max}(\omega)$ the local expression of the constraint for each frequency.

4. Oscillatory Source Amplitude

A first possibility to apply the strategy is to define by empirical observation a relationship between phase shift values and ground oscillation amplitude in interval $[0, \Omega_{max}] \times [0,\pi]$, ie to evaluate the function $A_G(\omega,\Delta\varphi)$, and to construct a set of simple fuzzy rules of car-driving type guaranteeing the satisfaction of amplitude condition in eqn(9). The procedure may be tedious as it has to be set up each time, and it is more convenient to construct directly the function $A_G(\omega,\Delta\varphi)$. This in turn splits into two problems : to calculate the oscillatory source at the piled object from eqns(5), and to determine the transmitted amplitude at interesting location from resolution of oscillation ground propagation. This last problem has already been investigated elsewhere[4], and only the first one will be analysed in the sequel from eqns(5). As mentioned earlier, this implies solving these equations with U as a source term and to use eqns(6,7,8) to determine the real input control $u(t)$ producing the source U. With $\Delta = Z - z$ and letting $\Delta_s = Z_s - z_s$ given by $\omega_M^2 \Delta_s = -\omega_g^2 - f_a$, eqns(5) reduces to

$$\begin{aligned} &\ddot{\overline{\Delta}} + \overline{\omega}^2 \overline{\Delta} = H + f_s \, \mathrm{sgn}(\dot{Z} - \dot{\Delta}) + U(t) \\ &\ddot{Z} = -\omega_M^2 \overline{\Delta} \end{aligned} \qquad (10)$$

with $\overline{\omega} = (\omega_M^2 + \omega_m^2)^{1/2}$, $\overline{\Delta} = \Delta - \Delta_s$,

$$H = [1 + (\omega_m/\omega_M)^2]\,\omega_g^2 + (\omega_m/\omega_M)^2 f_a - f_p \qquad (11)$$

and (formal) solution

$$\Delta(t) = \Delta_s - \frac{H \mp f_s}{\overline{\omega}^2} + \overline{\Delta}_0 \Delta_1(t) + \dot{\overline{\Delta}}_0 \Delta_2(t) + \int_0^t \Delta_2(t - t')U(t')dt' \qquad (12)$$

where $\overline{\Delta}_0 = \Delta_0 - \Delta_s$, $\dot{\overline{\Delta}}_0 = \dot{\Delta}_0$, $\Delta_1(t) = \cos \overline{\omega}\, t$, $\Delta_2(t) = (\sin \overline{\omega}\, t)/\ \overline{\omega}$. One then gets

$$Z(t) = P^{\pm}(t) - \tilde{Z}(t) \equiv Z_0 + \dot{Z}_0 t - \frac{H \mp f_s}{2\overline{\omega}^2} t^2 - \tilde{Z}(t) \qquad (13)$$

$$\tilde{Z}(t) = \overline{\Delta}_0 \Delta_1(t) + \dot{\overline{\Delta}}_0 \Delta_2(t) - \int_0^t dt' \int_0^{t'} U(t'')dt'' + \int_0^t \Delta_2(t - t')U(t')dt' \qquad (14)$$

The expression of $Z(t)$ contains two different parts, the parabolic time varying one $P^{\pm}(t)$, depending on the sign, and the oscillatory one $\tilde{Z}(t)$. To improve the penetration into the ground, the sign of the coefficient of t^2 should be always negative and as large as possible in absolute value. This means from eqn(11) that the applied force f_a should be above a critical value

$$f^{+}_{a,crit} = \frac{\omega_M^2}{\omega_m^2}\left\{ f_p - \left[1 + \frac{\omega_m^2}{\omega_M^2}\right]\omega_g^2 + f_s \right\} \qquad (15)$$

easier to reach as the ground reaction f_p is smaller, ie as vibratory unit frequency is larger. Then the trajectory goes monotonically downward along either of the two parabolas $P^{\pm}(t)$ defined in eqn(13), and goes up and down if $f^{-}_{a,crit} < f_a < f^{+}_{a,crit}$.

The other term $\tilde{Z}(t)$ fixes the amplitude and the frequencies of ground vibration oscillations, and depends itself on two terms, the mechanical mass oscillation with frequency $\overline{\omega}$ and the source term from input $U(t)$ which combines its own frequencies with $\overline{\omega}$. Its amplitude is typically given by

$$\left|\tilde{Z}(t)\right| = \left[\overline{\Delta}_0^2 + \dot{\overline{\Delta}}_0^2 + |U|^2\right]^{1/2} \qquad (16)$$

5. Determination of Source Term From Control Input

As indicated earlier, there remains to determine the apparent input to mechanical part $U(t)$ from its real input control source $u(t)$. This would rest upon solving the remaining system of eqns(6,7,8), but what is only needed is to reconstitute the input $u(t)$ which delivers the output $U(t)$ from eqns(6,7,8). A classical weaker form in control theory is to determine $u(t)$ in state space representation so that the output $U(t)$ asymptotically tracks a pre-assigned function $U_d(t)$ dictated by the research of smooth, monotonic and fast enough time response of the angular variables of the vibratory units[5]. Here elimination of the pressure variables p_j^+ and p_j^- from the subsystem with fixed j leads to the final general equation

$$\dot{Z}_j + 2A_j(\varphi_j, \varphi_1 - \varphi_2) - \frac{2\,\dot{\varphi}_j\, Z_j}{Z_j - B_j(\varphi_j)} = 0 \qquad (17)$$

with $Z_j = \lambda A_j^2(\varphi_j)/u_j^2$, $\beta_j = \mu_j \omega_g^2/\lambda\theta$, $B_j = \overline{p}_j - \beta_j \cos\varphi_j - \ddot{\varphi}_j/\lambda\theta$, $\overline{p}_j = (p_{sj} - p_{tj})/\lambda$ and

$$A_j(\varphi_j, \varphi_1 - \varphi_2) = \ddot{\varphi}_j/\lambda\theta + +2\dot{\varphi}_j - \beta_j \sin\varphi_j.\dot{\varphi}_j + (-)^{j-1}\overline{\omega}_j^2[\Gamma(\varphi_1 - \varphi_2 - \pi/2) + \Gamma(-\pi/2 - \varphi_1 + \varphi_2)](\dot{\varphi}_1 - \dot{\varphi}_2) \qquad (18)$$

where $\overline{\omega}_j^2 = \omega_j^2/\lambda\theta$, $j = 1,2$ and $\Gamma(x) = 1$ for $x>0$ and 0 for $x<0$. Eqn(17) determines directly the control input $u_j(t)$ which produces the output $\varphi_j(t)$. So choosing $\varphi_j(t)$ in previous class, it is possible to verify that the corresponding $u_j(t)$ is doable. As an example one could first consider stationary situation where both angular variables are rotating with same frequency Ω_0 and have a constant phase difference $\Delta\varphi_0$. If furthermore $\beta_j << \mathrm{Min}[2, \bar{p}_j]$, A_j and B_j simplify to $A_j = 2\Omega_0$ and $B_j = \bar{p}_{j\ t}$ so that eqns(17) decouple and now become

$$\frac{dY_j}{d\tau} + 2 - \frac{Y_j}{Y_j - \bar{p}_j} = 0 \qquad (19)$$

with $Y_j = 4\lambda\Omega_0^2 X_j$, $\tau = 2\Omega_0 t$. Its explicit solution is given by

$$\frac{\tau}{\bar{p}_j} = \Phi(\chi_j) \equiv \frac{1}{\chi_{j0}^2} - \frac{1}{\chi_j^{\ 2}} - \log\frac{1-2\chi_j^{\ 2}}{1-2\chi_{j0}^2} \qquad (20)$$

from which the input control is $u_{j,st} = 2\lambda\Omega_0\,(\lambda/\pi)^{1/2}\chi_j(\tau)$ with $\chi_{0j} \leq 1/2$ and $\chi_j(\tau) = \Phi^{-1}(\tau/\bar{p}_j)$ from eqn(20). It represents a monotonically growing curve from χ_{0j} to $1/2$ when the time runs from 0 to ∞. So it is possible to maintain single frequency oscillation Ω_0 of vibrating unit for long time with this smooth control input provided the final (highest) value is reachable. It should be noticed that exactly the same equation (17) is found with now $Y_j = \lambda\dot{\varphi}_j^2 X_j$ and $\tau = \varphi_j(t)$ in the more general case where $\varphi_j(t)$ are arbitrary but satisfy the inequality $\ddot{\varphi}_j / \dot{\varphi}_j << \lambda\theta$ and the condition $\varphi_1(t) - \varphi_2(t) \in [-\pi/2, \pi/2]$. To move the phase shift from initial value $\Delta\varphi_{0j}$ to final one $\Delta\varphi_{fj}$ at time T in this case, one should just program the control inputs $u_j(t)$ so that they follow the same curve $\Phi^{-1}(\varphi_j(t))$ from eqn(20) but with a different timing fixed by $\varphi_j(t)$. For instance one possible choice is to take

$$\varphi_1(t) = \Omega_0 t \ ; \ \varphi_2(t) = \varphi_1(t) + \Delta\varphi(t) \qquad (21)$$

where $\Delta\varphi(t)$ is a smooth monotonic S-type function such that $\Delta\varphi(0) = \Delta\varphi_0$ and $\Delta\varphi(T) = \Delta\varphi_f$ where $\Delta\varphi_0$, $\Delta\varphi_f \in [-\pi/2, \pi/2]$. To properly scale the input, it will only be necessary to verify that its final required power level from normalising expression above is effectively available. With this type of functions it is possible to monitor the phase shift between the angular variables of vibratory units and to ultimately control the amplitude of the produced oscillation from eqn(16) with the use of eqn(8). One then gets with eqns(21)

$$U(\tau) = \Omega_0\left[\mu_1^2 + \mu_2^2 + 2\mu_1\mu_2\cos\Delta\varphi(\tau)\right]^{1/2} \qquad (22)$$

from which it is possible to satisfy eqn(9) by proper programming of $\Delta\varphi(t)$ in operation time interval. Finally, to satisfy the smoothness requirement $\ddot{\varphi}_j / \dot{\varphi}_j << \lambda\theta$ for which the analysis applies (here $0 << \lambda\theta$), one possible way is to raise the rotation of eccentric masses up to high enough value with $\Delta\varphi(t) = \pi$ before starting to pile, and to monitor after the phases as in eqns(21).

6. Conclusions

To determine the level of ground oscillation generated by vibrating piling units used in urban earth and building works, the dynamics of the complete system are needed in order to account for all elements of the chain from power source to observation point. So the equations describing these dynamics which include both the motion of eccentric masses generating the vibrations and the motion of piled object with all their interactions have been set first. They can be split into two subsystem concerning the piled object dynamics with a fictitious input created by the vibrating unit, and the vibrating unit dynamics themselves from which it should be verified that the fictitious input is realisable with input from real power source. It has been possible to solve these two subsystems in such a way that useful properties can be obtained. From the first one, the piled object trajectory is obtained as the sum of a smooth monotonic parabolic time depending motion and an oscillatory one with combination of mass and vibrating unit base frequencies. The first motion is possible if the constant pressing force is larger than an explicit threshold value expressed in terms of system parameters. The amplitude of the second oscillatory motion is also evaluated in terms of the fictitious input source, so that there remains from the second subsystem to calculate it in terms of the real control input. This has been explicitly done when assuming that the normalised mass ratio between the piled object and the eccentric vibrating unit is small enough, in which case the time dependence of the control input is expressed in term of the phase angle of eccentric rotating mass. For regular and smooth enough time functions, the behaviour of control input is itself a regular bounded and self-similar one in the sense that it is the same function of the phase angle, so it can be pre-programmed once for all.

Appendix

The analytic solution of eqns(5,6,7,8) is also obtained in the (realistic) case where the vibratory units frequency Ω_1, Ω_2 are very large compared to mechanical characteristic frequencies ω_m, ω_M, ω_g associated to mass displacements. In this case one can write $U = U(t/\varepsilon)$ and use appropriate formalism for resolution. With $X = (Z, \dot{Z}, z, \dot{z})$, eqns(5) take the form

$$\frac{dX}{dt} = F\left(X, t, \frac{t}{\varepsilon}\right) \qquad \text{(A1)}$$

with

$$F = col\begin{pmatrix} X_2, \omega_M^2 (X_3 - X_1) - \omega_g^2 - f_a, X_4, \\ \omega_M^2 (X_1 - X_3) - \omega_g^2 - f_s \,\mathrm{sgn}(X_4) + U(t/\varepsilon) \end{pmatrix}$$

Splitting now in its slow and fast component $X = [X] + \{X\}$, and observing that integral on fast time writes $< F > = \int F(.,.,t/\varepsilon)\, dt = \varepsilon \int F(.,.,u)\, du$, it is possible to develop eqn(10) order by order in ε. One then gets to first order

$$\frac{d[X]}{dt} = \left[F\left([X], t, \frac{t}{\varepsilon}\right)\right] \qquad \text{(A2)}$$

$$\{X\} = < \left\{ F\left([X], t, \frac{t}{\varepsilon}\right)\right\} > \qquad \text{(A3)}$$

and to second one

$$\frac{d[X]}{dt} = \left[F\left([X], t, \frac{t}{\varepsilon}\right)\right] + \left[< \left\{ F\left([X], t, \frac{t}{\varepsilon}\right)\right\} > .\nabla F\left([X], t, \frac{t}{\varepsilon}\right)\right] \qquad \text{(A4)}$$

$$\{X\} = < \left\{ F\left([X], t, \frac{t}{\varepsilon}\right)\right\} > + < \left\{ < \left\{ F\left([X], t, \frac{t}{\varepsilon}\right)\right\} > .\nabla F\left([X], t, \frac{t}{\varepsilon}\right)\right\} > \qquad \text{(A5)}$$

which gives explicit expression for the solution Z, z when reporting $F(X,t,t/\varepsilon)$.

References

[1] Reusch, D. (2000) Automatic controlled vibratory pile driving with shock stress limitation. Proceedings of 17th International Symposium on Automation and Robotics in Construction, Taipei, Taiwan , Sept. 18-20.

[2] Launis, S., Keskinen, E.K. Raunisto, Y. (1999) Dynamic Analysis of Pile Penetration During Vibratory Driving - A Comparative Study, of Linear and Nonlinear Models, Proceedings of the IffToMM 10th World Congress on the Theory of Machines and Mechanisms, Oulu, Finland, June 20-24, 1631-1637.

[3] Keskinen E., Kivinen J.M., Miettinen J., Cotsaftis M., Hildebrand R., Raunisto Y. (2003) Dynamics and monitoring of vibratory piling process, Proceedings of Intern. Symposium on Robotics and Automation in Construction, Eindhoven, the Netherlands, Sept 21-25.

[4] Keskinen E., Launis S., Cotsaftis M., Raunisto Y. (2001) Ground Vibrations Dynamics during Pile Hammering, Proc. 6th Intern. Conf. on Integral Methods in Science and Engineering (ISME), June 12-15, Banff, Alberta, Canada

[5] Liu C. (1983) General Decoupling Theory of Multivariable Process Control Systems, Lecture Notes in Control and Information Theory, n°53, Springer-Verlag, New-York; Bocker J., Hartmann I., Zwanzig C. (1986) Nonlinear and Adaptive Control Systems, Springer-Verlag, New-York

Rapid 3D Object Modeling Using 3D Data from Flash LADAR for Automated Construction Equipment Operations

Hyojoo Son[1], Changwan Kim[2] and Yoora Park[3]

[1] Research Assistant, Dept. of Architectural Engineering, Chung-Aug University, 221 Hueksuk-dong, Dongjak-gu, Seoul, Korea 156-756, PH (82) 2-825-5726; FAX (82) 2-825-5726; e-mail: hjson0908@wm.cau.ac.kr
[2] Assistant Professor, Dept. of Architectural Engineering, Chung-Aug University, 221 Hueksuk-dong, Dongjak-gu, Seoul, Korea 156-756, PH (82) 2-820-5726; FAX (82) 2-812-4150; e-mail: changwan@cau.ac.kr
[3] Research Assistant, Dept. of Architectural Engineering, Chung-Aug University, 221 Hueksuk-dong, Dongjak-gu, Seoul, Korea 156-756, PH (82) 2-825-5726; FAX (82) 2-825-5726; e-mail: saeara10@nate.com

Abstract

Automated recognition and modeling of 3D objects located in a construction work environment that are difficult to characterize or are constantly changing is critical for autonomous heavy equipment operation. Such automation allows for accurate, efficient, and autonomous operation of heavy equipment in a broad range of construction tasks by providing interactive background information. This paper presents 3D object recognition and modeling system from range data obtained from flash LADAR, with the goal of rapid and effective representation of the construction workspace. The proposed system consists of four steps: data acquisition, pre-processing, object segmentation on range images, and 3D model generation. During the object segmentation process, the split-and-merge algorithm, which separates a set of objects in a range image into individual objects, is applied to range images for the segmentation of objects. The whole process is automatic and is performed in nearly real time with an acceptable level of accuracy. The system was validated in outdoor experiments, and the results show that the proposed 3D object recognition and modeling system achieves a good balance between speed and accuracy, and hence could be used to enhance efficiency and productivity in the autonomous operation of heavy equipment.

Keywords: 3D object modeling; 3D object recognition; construction automation, construction heavy-equipment, flash LADAR, range image segmentation

1. Introduction

In recent years, with fundamental advances in sensor technology, it is becoming ever more feasible to automate construction equipment in ways that would help improve efficiency and safety of equipment operations on construction sites. While automation of heavy equipment has the potential to make an important contribution to productivity improvement, what is needed is an efficient way to represent a workspace in 3D and incorporate that representation into control of equipment operations. There are requirements that 3D modeling methods for use in construction automation have to satisfy; data acquisition speed, level of intricacy, versatility, and automated data processing.

Recently, 3D modeling methods have been investigated to represent construction workspace for several applications such as as-built drawings, visual feedback to equipment operators, and construction materials tracking. Most research on 3D modeling in the construction industry have employed large and expensive 3D laser scanner to produce dense point clouds. While 3D laser scanning can produce very detailed models of the scanned scenes, which are useful for obtaining as-built drawings of existing structures, extensive processing of the received point clouds is needed in order to construct the 3D model, thereby processing the entire 3D scanning process too laborious and time-consuming for the intended applications. The burdens imposed by a 3D laser scanner in terms of processing time generally preclude the use such technology for real-time decision-making.

An alternative method is based on the use of flash LADAR which encompasses a new generation of scannerless LADAR devices. A flash LADAR device produces an image of the observed scene in which each pixel value represents the intensity and range of the corresponding image area (Uijt de Haag et al. 2008). When compared to a 3D laser scanner, a scannerless flash LADAR device is smaller in size, lighter in weight, and lower in cost and has advantage of acquiring data in real-time (Stone and Juberts 2002; Habbit et al. 2003). With the capability of capturing range data at high speed, flash LADAR produces a 3D image an entire scene in a single data acquisition step; moreover, it can capture moving objects, and hence can provide both static and dynamic information (Habbit et al. 2003). As a result, flash LADAR is beneficial for real-time applications such as obstacle detection, equipment navigation and object recognition (Price et al. 2007). Although flash LADAR has brought a new means of achieving real-time 3D model generation from range data, few research studies on this technology have been undertaken in the construction industry (Teizer et al. 2007; Gong and Caldas 2008; Kim et al. 2008). Much more work needs to be done in order to achieve automated recognition and modeling of objects on a construction site from range data and reconstruction of useful 3D model in near-real-time.

The aim of this research is to develop a system for 3D object recognition and modeling using range data from flash LADAR, with the goal of rapidly and effectively representing construction workspaces. To achieve this purpose, a system is proposed, consisting of algorithms that recognize the objects in the scene, together with methods for automatically extracting feature points related to those objects and generating bounded 3D models of each object. And outdoor experiments have been performed to test the performance of the proposed method.

2. Framework for 3D Object Modeling for Use in Heavy Equipment Operation

In this section, an overview of the framework for the proposed automatic 3D object recognition and modeling system using range image is presented. The process used in the 3D object recognition and modeling system proposed herein is outlined in the flowchart in Figure 1.

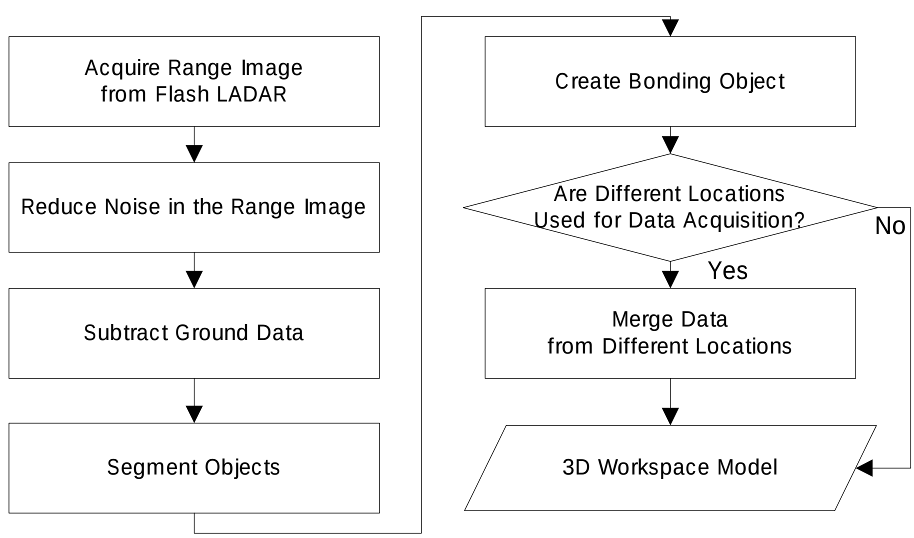


Figure 1 Proposed Process for 3D Workspace Modeling

The first step of the 3D object recognition and modeling process is to acquire the 3D data that adequately covers the 3D scene. As mentioned earlier, workspace models are required to express dynamic work environment of a construction site effectively and in near-real-time for construction automation applications (Kim et al. 2006). For this reason, not only high frequency of updated of 3D data acquisition but also an acceptable level of accuracy are needed for reliable and successful 3D object modeling. Thus, in this study, the 3D data of the objects were acquired by using a SwissRanger SR-3000 flash LADAR, since it provides the most adapted trade-off between the data acquisition speed and the accuracy of the 3D data in case of real-time applications (Bosche et al. 2005).

Once the range image acquired, the range image acquired from high speed range scanner such as flash LADAR contains noise of considerable level (Frome et al. 2004). If the noise is not reduced, it may affect a negative effect on object recognition; therefore, pre-processing is needed. For this, data filtering method is performed to reduce the noise influence. In addition, objects on the ground are one that its boundary between object region and ground region is hard to recognize. To detect and remove ground data, ground subtraction technique is used.

Pre-processing step is completed; the feature points of objects are extracted without distinction between different objects. The range image should be separated into individual objects. This is the segmentation process which is required to recognize the objects in a scene. In this study, a split-and-merge algorithm is applied to separate the set of extracted feature points into individual objects

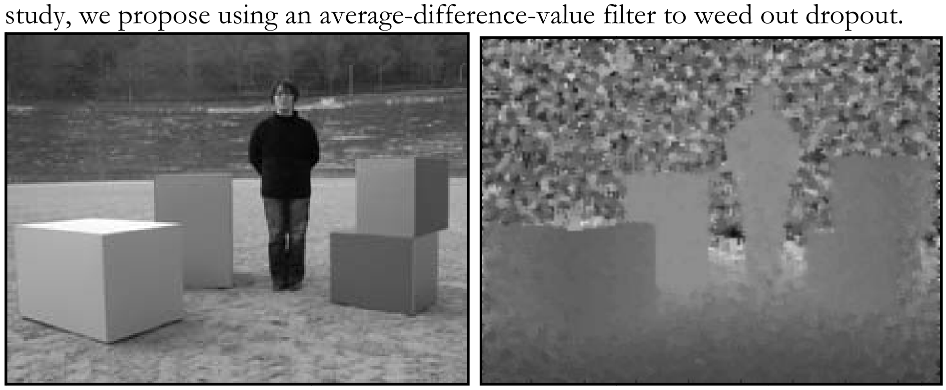
Then bounding models of objects representing various construction site scenes are created through the use of a general class of geometric primitives. In cases where the operator's view from any one reference point is limited and the data have to be acquired from two or more locations, the data sets obtained from the different locations are merged into a single data set having one common coordinate system by applying an ICP (iterative closest point) algorithm. The 3D models generated from data acquired on-site via the process described above are useful as a tool for providing interactive feedback to equipment operators. Additionally, the models can be shown as 3D simulations that offer equipment operators an opportunity to experience the results of certain aspects of the tasks at hand before actual operation.

3. Experiments and Results

This section describes the detailed methods of 3D object recognition and modeling system with the discussion of outdoor experimental results. Outdoor experiments were conducted to establish validation for the proposed object recognition and modeling system.

3.1. Pre-processing

After the data acquisition using flash LADAR, 25,344 data points were acquired per range image. Range image acquired from flash LADAR in an outdoor environment tend to contain large regions of dropout, because of the measurement limits of flash LADAR and outdoor environmental conditions (Frome et al. 2004). Figure 2(b) shows a range image in which fluctuations in the level of gray indicate false range values. Thus, in this study, we propose using an average-difference-value filter to weed out dropout.



(a) (b)

Figure 2 (a) Photographic Image, (b) Range Image

The average difference value ADV employed in this research is the average of the differences between the value of a given pixel and those of its eight neighbors in the 3×3 window centered at that pixel. If the value of ADV is larger than some predefined threshold, the central pixel is assumed to be corrupted by dropout and is eliminated. Otherwise, the central pixel is left unchanged. Throughout the process, points with a range above a threshold value of 0.6 were weeded out. This threshold was selected after performing a set of experiments and finding that it successfully detected the noisy regions of range images acquired from flash LADAR.

After the results of the average-difference filtering were applied, there was still speckle noise left in the image, especially in the object region, which caused measurement errors to creep in. In this study, a 3×3 median filter was used to remove speckle noise and render the surfaces of objects more uniformly. A median filter is useful in eliminating speckle noise in a range image while preserving edge information (Doss 2004). As shown in Figure 3(a), noise was reduced to an acceptable level after using the proposed filtering method, while object regions were preserved.

However, in this case, there is the large part of the points corresponding to the ground floor which makes difficult to segment the objects. Thus, the next step undertaken in the object modeling is removal of data corresponding to the ground floor of the scene. In cases where the sensor is located on a rigid body, and the height and vertical angles are known, ground plane detection is pretty straightforward. After the data are subjected to an appropriate coordinate rotation, the ground data can be extracted using a threshold value (Bostelman et al. 2006). The result of such an approach used for removal of ground data in this research is shown in Figure 3(b). After the pre-processing step, about the 65% (16,431 points) of the raw data points were filtered and removed.

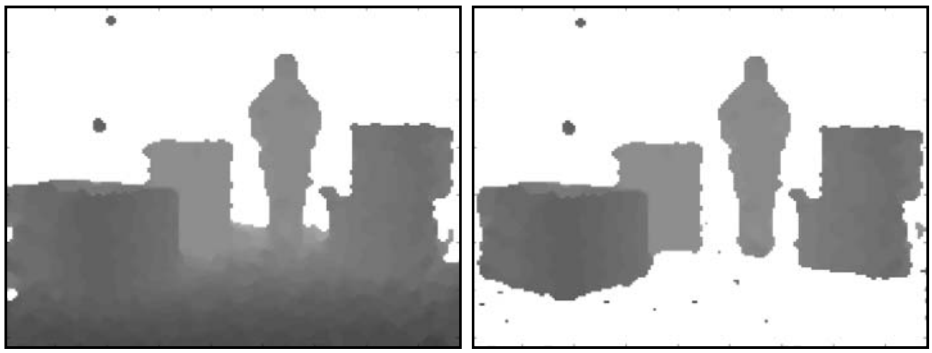


Figure 3 (a) Result of Data Filtering Process, (b) Filtering Out of Ground Floor Related Data Points

3.2. Object Segmentation

In the pre-processing stage, range values for the unwanted regions are effectively extracted. Once that is accomplished, a process of decomposition of the resulting image into separated objects is undertaken. Various methods for object segmentation, which are classified as boundary-based approaches and region-based approaches, have been developed (Xiang and Wang 2004). Boundary-based approaches identify the edges based on discontinuities (pixel differences) and link up edges to produce closed boundaries for individual objects (Lin and Talbot 2001). However, their applicability is limited by the fact that it is difficult to find complete boundary information for an object, especially in a noisy image (Ikeuchi 1987). Region-based approaches take a noisy image and distinguish coherent regions that satisfy a predefined homogeneity condition, and then use those coherent regions to identify the objects of interest (Kelkar and Gupta 2008). In this study, split-and-merge segmentation, which is one of the main methods used in region-based approaches, was adopted (Lin and Talbot 2001). This method has the advantage of simplicity of use as well as computational efficiency by a combination of splitting and merging of regions in the image (Salih and Ramli 2001; Sun and Du 2004).

The object segmentation process based on split-and-merge algorithm comprises following two steps: splitting and merging. At the first split operation, the splitting process starts with the complete range image Figure 3(b) as a single region R . If R is inhomogeneous, it is split into four subregions—in particular, four rectangular blocks of equal size. This process is repeated recursively until all subregions are homogeneous. Since the splitting process might have split up homogeneous regions, a merging process is then used to test the homogeneity of adjacent regions and merge them into a single region if their union is homogeneous. Through the object segmentation process, four objects were successfully separated as shown in Figure 4.

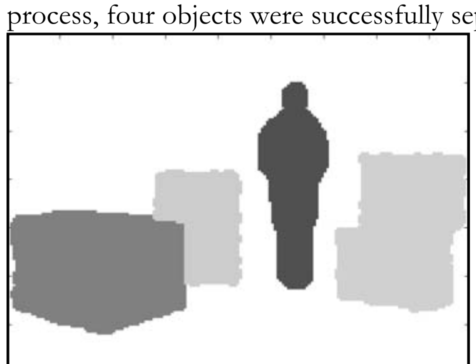


Figure 4 Result of the Object Segmentation Process

3.3. Model Generation

In the object segmentation process, the feature points of objects were successfully segmented and identified into four objects. Based on the segmented feature points of each object, each object points were connected to generate 3D model by using a convex hull algorithm as shown in Figure 5. The convex hull algorithm used in this research is an incremental algorithm developed by Barber et al. (1996); it is well suited for rapid 3D object modeling, because its use in the generation of bounding models to represent the spatial shapes of 3D objects is compact, fast, and relatively efficient (Kim et al. 2006).

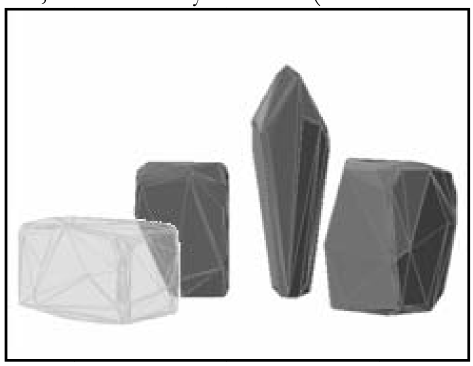


Figure 5 3D Model Generated from One Point of View (the object at left is shown as a translucent model)

In the case at hand, there were two objects in the left portion of the scene, one of which was partially occluded by the other, as shown in Figure 5. Also, in some cases, it may be impossible to acquire an image of the complete scene due to limitations of the sensor's field of view or for geometric reasons (Neugebauer 1997; Mure-Dubois and Hugli 2008). Thus, the supplemental 3D data was acquired from different location and segmented into separated objects. And data sets acquired from different locations were registered together with the respect to the same coordinate system. In this study, an ICP algorithm developed by Besel and McKay (Besl and McKay 1992) was used for that purpose. The ICP algorithm utilized here, which was designed for use with free-form surfaces, works by automatically matching the closest point in one set of 3D data to another set of 3D data (Allen et al. 2003). Through this process, complete 3D models were generated by making up for range image (for the second object in the left portion of the range) that were not acquired in the first round of data acquisition, as shown in Figure 6.

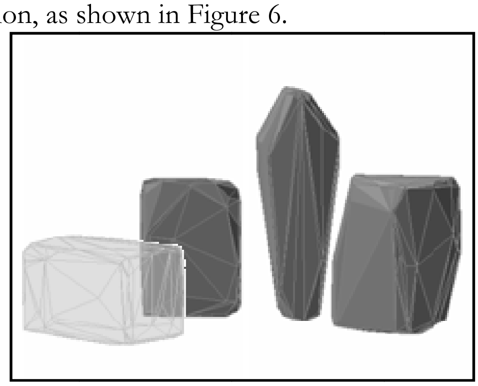


Figure 6 Merged 3D Models (the object at left is shown as a translucent model)

5. Conclusions

This study presented rapid and efficient 3D object recognition and modeling system for use in automated operation of construction equipment. Flash LADAR is shown to be a viable means of acquiring real-time spatial information in the form of a range image of a construction-site scene. The proposed data processing scheme, which includes pre-processing, object segmentation, and model generation, utilizes various algorithms to recognize the objects in the scene from the acquired data and then automatically generate a 3D model. The effectiveness of the proposed object recognition and modeling system was validated. Outdoor experimental results demonstrated that it takes only a few seconds to generate 3D models with the proposed method, hence that it can be used for automated object recognition and modeling of construction objects in near-real-time. Adoption of such a method could facilitate the safe and efficient operation of heavy equipment, since it not only provides a 3D graphical representation of the scene but also enables spatial analysis of the current set of conditions on a construction site.

Although the proposed modeling method has been applied in this research to the modeling of static objects only, we see no reason why it could not be extended to the modeling of dynamic objects. Also, the proposed method would seem to lend itself to use in control systems for heavy construction equipment, where it could be used, for example, as part of an obstacle avoidance system.

6. Acknowledgements

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (KRF-2008-313-D01164).

References

[1] Allen, P. K., Troccoli, A., Smith, B., Murray, S., Stamos, I., and Leordeanu, M. (2003). “New methods for digital modeling of historic sites.” IEEE Comp. Graph. Appl., 23(6), 32–41.

[2] Barber, B., Cobkin, D., and Huhdanpaa, H. (1996). “Quickhull algorithm for convex hull.” ACM Trans. Math. Software, 22(4), 469–483.

[3] Besl P. and McKay N. (1992). “A method for registration of 3-D shapes.” IEEE Trans. Pat. Anal. and Mach. Intel., 14(2), 239–256.

[4] Bosche, F., Haas, C. T., and Caldas, C. H. (2005). “3D CAD drawing as a priori knowledge for machine vision in construction.” Proc., 1st Ann. Inter-University Symp. on Infrastructure Management (AISIM), Waterloo, Canada.

[5] Bostelman, R., Russo, P., Albus, J., Hong, T., and Madhavan, R. (2006). “Applications of a 3D range camera towards healthcare mobility aids.” Proc., 2006 IEEE Int. Conf. on Networking, Sensing, and Control (ICNSC ’06), Ft. Lauderdale, Fl., 416–421.

[6] Frome A., Huber, D., Kolluri, R., Bulow, T., and Malik, J. (2004). “Recognizing objects in range data using regional point descriptors.” Proc., 8th European Conf. Computer Vision (ECCV), Prague, Czech Republic, 224–237.

[7] Doss, N. (2004). “3D modeling of biological structures.” BS thesis, Univ. Western Australia, Nedlands, WA 6907, Australia.

[8] Gong, J., and Caldas, C. H. (2008). “Data processing for real-time construction site spatial modeling.” J. Autom. Constr., 17(5), 526–535.

[9] Habbit, R. D., Nellums, R. O., Niese, A. D., and Rodriguez, J. L. (2003). “Utilization of flash ladar for cooperative and uncooperative rendezvous and capture.” Proc., SPIE, 5088, Orlando, Fl., 146–157.

[10] Ikeuchi, K. (1987). “Generating an interpretation tree from a CAD model for 3D-object recognition in bin-picking tasks.” Int. J. Comput. Vision, 1(2), 145–165.

[11] Kelkar, D., and Gupta, S. (2008). “Improved quadtree method for split merge image segmentation.” Proc., 2008 1st Int. Conf. on Emerging Trends in Engineering and Technology, Nagpur, India, 44–47.

[12] Kim, C., Haas, C. T., Liapi, K. A., and Caldas, C. H. (2006). “Human-assisted obstacle avoidance system using 3D workspace modeling for construction equipment operation.” J. Comput. Civ. Eng., 20(3), 177–186.

[13] Kim, C., Son, H., Kim, H., and Han, S. (2008), “Applicability of flash laser distance and ranging to three-dimensional spatial information acquisition and modeling on a construction site.” Can. J. Civ. Eng., 35(11), 1331–1341.

[14] Lin, Z., Jin, J., and Talbot, H. (2001). “Unseeded region growing for 3D image segmentation.” Proc., Selected papers from Pan-Sydney Area Worksh. on Visual Information Processing (VIP2000), Sydney, Australia, 31–37.

[15] Mure-Dubois, J., and Hugli, H. (2008). “Fusion of time of flight camera point clouds.” Proc. IEEE Worksh. on Multi-camera and Multi-model Sensor Fusion Algorithms and Applications (M2SFA2), Marseille, France.

[16] Neugebauer, P. J. (1997). “Geometrical cloning of 3D objects via simultaneous registration of multiple range images.” Proc., 1997 Int. Conf. Shape Modeling and Applications (SMA ’97), 130–139.

[17] Price, M., Kenney, J., Eastman, R. D., and Hong, T. (2007). "Training and optimization of operating parameters for flash ladar cameras." Proc., 2007 IEEE Int. Conf. on Robotics and Automation, Roma, Italy, 3408–3413.

[18] Salih, Q. A., and Ramli, A. R. (2001). "Region based segmentation technique and algorithms for 3D image." Proc., Int. Symp. on Signal Processing and Its Applications (ISSPA), Kuala Lumpur, Malaysia, 747–748.

[19] Stone, W. C., and Juberts, M. (2002). "Towards the ultimate construction site sensor." Proc., 19th Int. Symp. on Automation and Robotics in Construction (ISARC), National Institute of Standards and Technology, Gaithersburg, Maryland, 393–400.

[20] Sun, D., and Du, C. (2004). "Segmentation of complex food images by stick growing and merging algorithm." J. Food Eng., 61(1), 17–26.

[21] Teizer, J., Caldas, C. H., and Haas, C. T. (2007). "Real-time three-dimensional occupancy grid modeling for the detection and tracking of construction resources." J. Construct. Eng. Manag., 133(11), 880–888.

[22] Uijt de Haag, M., Venable, D., and Soloviev, A. (2008). "Implementation of a flash-ladar aided inertial navigator." Proc., IEEE/ION Position, Location, and Navigation Symp., Monterey, Ca., 832–837.

[23] Xiang, R., and Wang, R. (2004). "Range image segmentation based on split-merge clustering." Proc., 17th Int. Conf. on Pattern Recognition, Cambridge, UK, 614–617.

Development of Automated Inspection Robot and Diagnosis Method for Tile Wall Separation by Wavelet Analysis

F. Inoue[1], S. Doi[1] , T. Okada[2] and Y. Ohta[2]

[1]Technical Research Institute, Obayashi Corporation, 4-640 Simokiyoto, Kiyose-shi, Tokyo, 204-8558, Japan, PH (+81)-42-495-1275, FAX (+81)-42-495-0940, inoue.fumihiro@obayashi.co.jp
[2]Department of Science and Engineering, Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo, 169-5678, Japan, PH (+81)-3-5286-3267, FAX (+81)-3-5286-3267, yutaka@waseda.jp

Abstract

This paper describes the development of an automated inspection robot for detecting tile deterioration and a new diagnostic method for determining its existence and extent. The robot moves quickly along a vertical wall and stops to detect a tile's inner condition using a hammering sound. Tile separation commonly comprises outer exfoliation where the tile separates from the mortar concrete and inner exfoliation where the space between the substrate and the mortar concrete deteriorates. In order to detect these two modes by a hammering sound, we focused attention on wavelet analysis, which enables us to analyze the frequency element of the sound waveform. By comparing the wavelet volume rate expressing the characteristics of tile deterioration, a method of visually distinguishing the two exfoliation modes was established. The automated robot and the diagnostics method were used to perform a fast and highly accurate inspection of a tile wall.

1. Introduction

Tiles made from materials that are durable, fireproof, water-proof and decorative are used for external finishes on exterior walls of architectural buildings. However, many of these finishing materials have deteriorated with age and changes of building use. Typical deterioration symptoms include tile separation, cracking and layer separation. If they are left too long, tiles can break and large areas can suddenly fall off, and in extreme cases fatal accidents can occur. To prevent these problems, the deteriorated states need to be accurately confirmed and an appropriate repair method determined. External tile walls have previously been inspected by skilled workers using a hammering sound. However, this has led to variations in diagnosis, and it is feared that judgment can be impaired when work has continued for a long time at a great height.

In the high economic growth period after 1975 in Japan, more than 100,000 buildings were constructed in metropolitan areas. Thirty years have already passed since that period, and symptoms of tile deterioration have begun to appear. The Ministry of Land, Infrastructure and Transportation revised The Building Standard Law of Japan in 2008 with regard to the safety performance of exterior tile walls, requiring safety inspections of all tile walls ten years after construction.

As a background to tile inspection technology, for about ten years we studied several diagnostic methods and developed several kinds of inspection robots, finally resulting in the present robot system. The robot has a very compact body and shows good inspection performance for tile deterioration (Inoue, F., and Ohta, Y. (2004)). However, for the diagnostics method, some parameters indicating characteristics of sound waveform for detecting tile exfoliation have been found experimentally by applying mathematical functions, and used to diagnose symptoms of tile deterioration. These methods were found to be useful for the simple deterioration mode of tile surface separation, but not for the compound tile deterioration of separation of inner concrete layer, which can induce a fatal accident. Furthermore, it was very difficult for a skilled worker depending on his senses to distinguish between separation of outer tiles and inner layer deterioration.

This paper introduces a developed inspection robot and actual applications, as well as new diagnostic methods for compound tile deterioration using wavelet analysis, and these effects are explained in detail.

2. Development of Automated Inspection Robot

2.1 Outline of the inspection system using an automated robot

Figure 1 shows a schematic picture of the tile inspection system using an automated robot instead of a skilled worker. The inspection robot moves along the tiled face of the building while hanging from two wires suspended from a hanger truck and its sensors check various tile deterioration modes. The robot can be moved vertically by winding and unwinding these wires, and horizontally by moving the hanger truck. In an actual inspection, the robot is stopped at each tile face for a short time, and the deterioration conditions of the tile are examined by analyzing the impact sound of a small hammer and by observing camera data, and the diagnosis result is instantly indicated on a computer panel on the ground. All the tile tests and test result recordings are executed by automatic operation, except for a few manual operations.

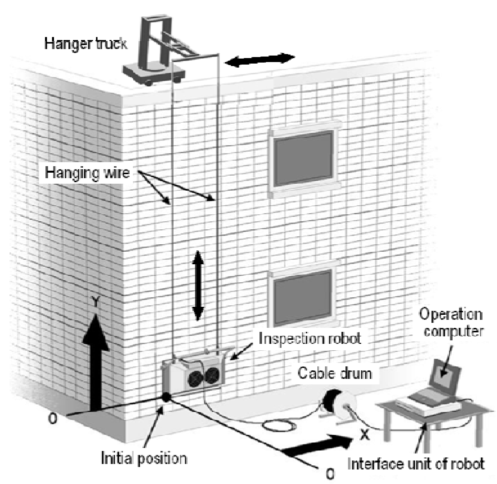


Figure1 Outline of the inspection system using an automated robot

2.2 Mechanism of Inspection Robot

Picture and Specification of the inspection robot is shown in Figure 2 and Table 1 respectively. The robot itself is composed of four parts; a wire winding part, vacuum fans, an inspection sensor part, and an operation and analysis part. By operating the wire winding part, the robot is situated vertically at a desired point. Two wires are rolled onto a rotating drum in the robot. The robot can climb up and down at a speed of 3-5 m/min in a height range of 40 m. The vacuum fans are used to press the robot body onto the tile face. The thrust force is more than 15 N, and the robot body is sufficiently stable against wind. The inspection part is composed of a mechanical hammer and a movement slide. The hammer can move to each tile position according to the tile size within a range of 450 mm. The hammer is electrically controlled by solenoid power and hits the tile face only once. The impact sound is measured by a small microphone installed near the point where the tile is hit. The total weight of the robot body is about 30 kg, and it can be operated easily even if the operator is not skilled or experienced in tile inspection.

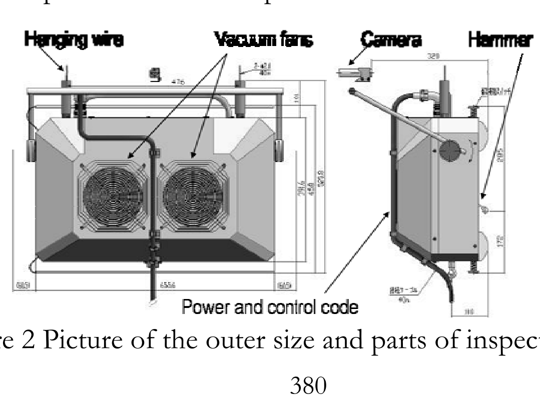


Figure 2 Picture of the outer size and parts of inspection robot

Table 1 Specification of the inspection robot

Inspection robot	Outer size	W780×L526×H250
	Weight	30 kg
	Power source	AC 100V
	Hanmmer Slide width	450 mm
	Climbing Speed	3-5 m/min.
	Thrast force	15 N
	Fan Power	100W ×2sets
	Measurement ability	450 m2/day
	Wire length	40 m
Hanger truck	Outer size	W1500×L1500×H500
	Weight	70 kg
	Operation	Manual
	Hanger top weigth	100 kg
Operation system	Drive interface	W310×L260×H50
	Controller	1 set

2.3 Analysis of Impact Sound and Conventional Diagnostic Methods

The inspection robot uses two analysis parameters to apply mathematical functions to express the characteristics of tile exfoliation (Inoue, F., and Ohta, Y. (2006)). The time from the point when an impact sound is measured to that when the value of each parameter value is calculated is very short, so the deterioration result can be diagnosed in real-time. The parameters are as follows.

(1) Peak rate value : C_f

This parameter is equal to the wave height value, and is the ratio of the maximum amplitude value X_{max} to the effective value of time waveform X_{rms}. It becomes higher when the tile is separated.

$$C_f = X_{\max} / X_{rms} \qquad (1)$$

(2) Frequency cross correlation value: $\rho_{x,y}$

This parameter is the frequency cross correlation between the sound waveform for each tile and that of a normal tile. It becomes lower when the correlation with a normal tile is decreased when the tile is separated. N is the number of data, X and Y are the sound waveform of the test tile and a normal tile, respectively, and μ and σ are the average and standard deviation of X and Y, respectively.

$$\rho_{x,y} = \left\{ \frac{1}{N} \sum_{i=1}^{N} (X_i - \mu_x) \cdot (Y_i - \mu_y) \right\} \Big/ \sigma_x \cdot \sigma_y \qquad (2)$$

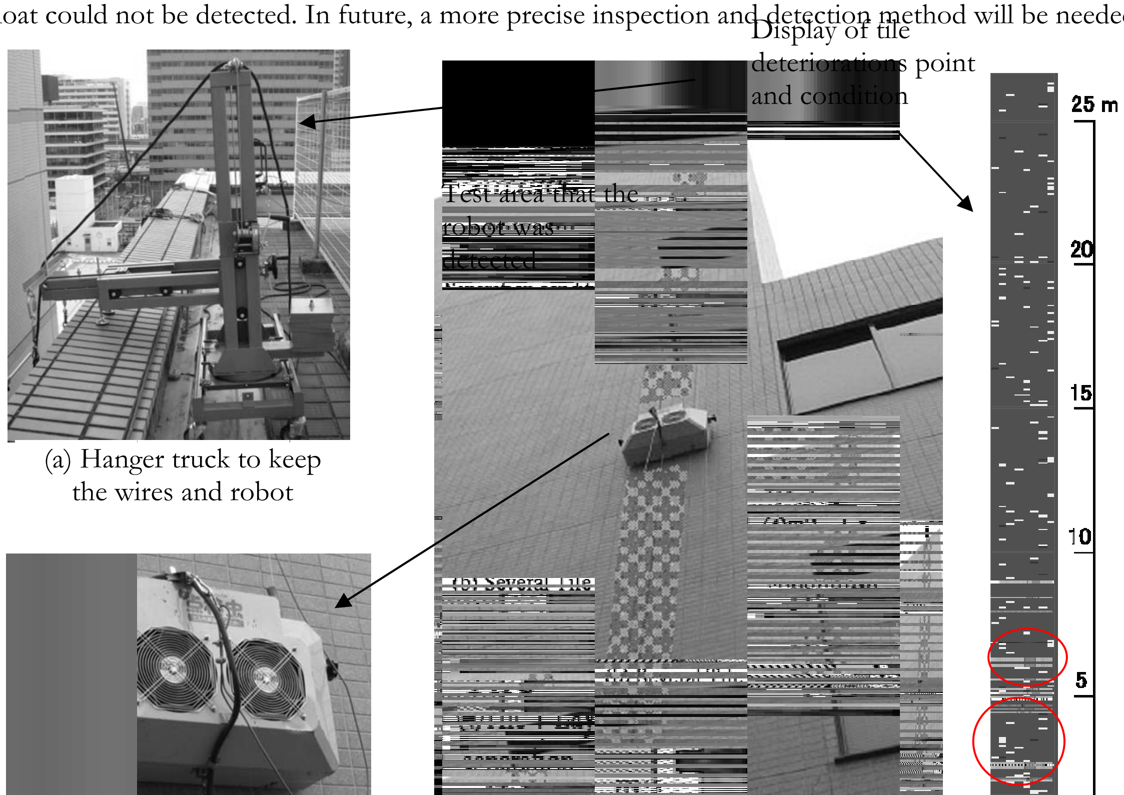
2.4 Application example of actual tile inspection and diagnostic result

An example of an actual tile inspection using the inspection robot is indicated in Fig.3. The tile wall, which was built 20 years ago, was about 25m high, and tile exfoliation and distortion of the wall was suspected. The robot was applied to shorten the work period and to accurately diagnose tile deterioration. At the actual site, the robot (Figure 3-(b)) was hung by wire from a hanger truck (Figure 3-(a)). The robot was moved up and down to continuously detect tile deterioration along a vertical line. In this case, the frequency cross correlation value $\rho_{x,y}$ was adopted as the diagnostic parameter.

The test range and the diagnostic result corresponding to the test point are indicated in Figure -(c). At the point that is short bar colored on the display, some abnormal tiles were detected. The red bar indicates that the value of the analysis parameter was lower than that of the yellow bar. In short, it is judged that tile deterioration corresponding to the red bar is very serious. Also, within some radius around the color bars, it is possible that a wide separation of the inner concrete layer including outer tiles may be generated. From the inspection result, the entire tile was quickly repaired and the danger of collapse of the outer wall was averted.

Up to now, tiles inspections of tens of buildings have been executed using this inspection robot and effective diagnostic results have been obtained. However, for the diagnostic parameters given in equations

(1) and (2), the deterioration state has not been determined in detail, or where or how separation has been generated. In particular, compound deterioration such as outward tail separation and inner concrete layer float could not be detected. In future, a more precise inspection and detection method will be needed.



(a) Hanger truck to keep the wires and robot

(b) Developed inspection robot

(c) Tile detection and indicating its diagnostic result

Figure 3 Actual tile inspection using the inspection robot and display the diagnostics result

3. New Diagnostic Method Using Wavelet Analysis

3.1 Characteristics of Tile around Deterioration for Finish Condition

The finish condition of a general tile wall is composed from the inside of the main concrete layer, the mortar layer, and the bonding tile layer, as shown in Figure 4. Many tile wall collapses are divided into the following two kinds of separations on the basis of an empirical rule.

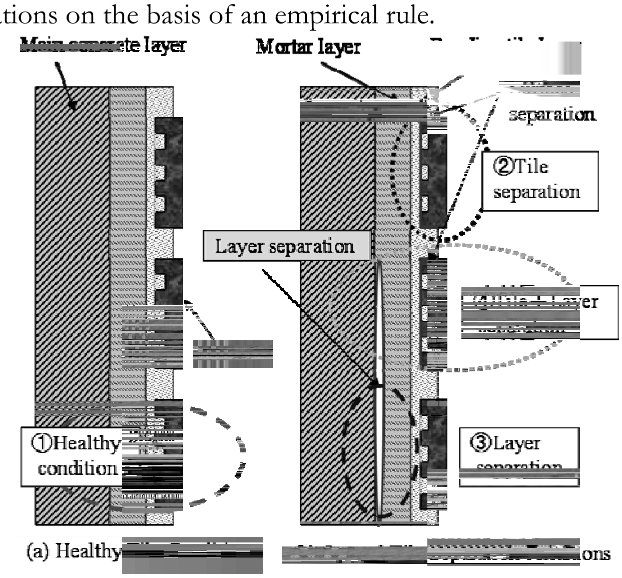


Figure 4 Patterns of the tile deterioration conditions

(1) Tile separation: the condition in which a tile space is generated by separation of the back or side of the tile from the bonding tile mortar.

(2) Layer separation: the condition in which a layer space is generated between the main concrete layer and the mortar layer, or between the mortar layer and the bonding tile layer.

Tile separation is a narrow deterioration that acts on the individual tile, but layer separation is a wide deterioration that acts on the thin layer, what we call "float condition", and then a large area of tile wall is likely to fall off the main concrete layer when the condition deteriorates. In an actual situation, tile separation and layer separation exist together in a complicated manner, and the condition around a tile is thought to be one of the following, as shown in Fig.4: ①Healthy condition, ②Tile separation only, ③Layer separation only, and ④Tile and Layer separation together. The next paragraph clarifies the method of diagnosing these four conditions.

3.2 Sound Waveform Analysis Method

Peculiar time waveforms corresponding to the tile condition were obtained as indicated in Fig.5 for the reflected sound when the tile is struck by the impact hammer. The rise of the impact time waveform measured at the normal tile (Figure 5-(a)) is lower and it decreases instantly. The rise of the impact time waveform measured at the deteriorated tile is higher and it decreases gradually (Figure 5-(b)), and the waveform frequency differs according to the deterioration condition. For a waveform having the characteristics of tile deterioration, wavelet analysis showing that the waveform can be transformed to the time and frequency range serially is considered to be suitable.

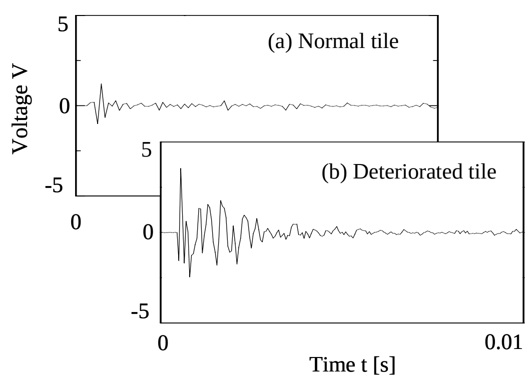


Figure 5 A peculiar time waveforms corresponding to the tile condition

3.3 Analysis by Wavelet Transformation

3.3.1 Introduction of wavelet transformation

The value of the wavelet transformation (W*f*) of time waveform *f*(t) to mother wavelet function φ(t) is given by equation (3) (Chui, C. K. (1992)) . Here, a and b express the constant scale number of the frequency and time axis for mother function φ(t) .

$$(W_\varphi f)(b,a) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{|a|}}\ \varphi\left(\frac{t-b}{a}\right)\ f(t)\ dt \qquad (3)$$

Using the wavelet transformation, it is very important what function is selected in the mother wavelet function φ(τ). Generally, analytical accuracy of the wavelet transformation is said to improve as the pattern and correlation of the measurement waveform and mother function improves. Therefore, an original mother function was designed using a new design approach explained in the next paragraph (Daubechies, I. (1992)).

3.3.2 Design of original mother function

The order transforming from the actual data to the mother function is indicated below. A sound waveform measured from the normal tile was applied as actual data (Figure 6-(a)).

(1) Detection of an initial time of the waveform: First, the time at which the absolute value of the waveform reaches a maximum value is selected as indicated at P (Figure 6-(b)), and going back from P, the time at which the sign changes for the first time is specified as the initial time at Q (Figure 6-(a)).

(2) Detection of a finishing time of the waveform: Calculating the individual auto-correlation of the waveform, the first point at which the value of auto-correlation data is less than 0.1 is specified as the finishing time (Figure 6-(c)).

(3) Arrangement of the waveform: As the mother function should generally satisfy an admissible condition, the symmetry of the trimmed waveform is ensured by the following approach (Figure 6-(d)).

- The reverse of the time order and the sign of waveform
- Combinations of the trimmed waveform and the original waveform

(4) Normalization of the arranged waveform: By calculating the norm value of the arranged waveform, the entire waveform is divided by the norm value. Such a normalized waveform is defined as an original mother function.

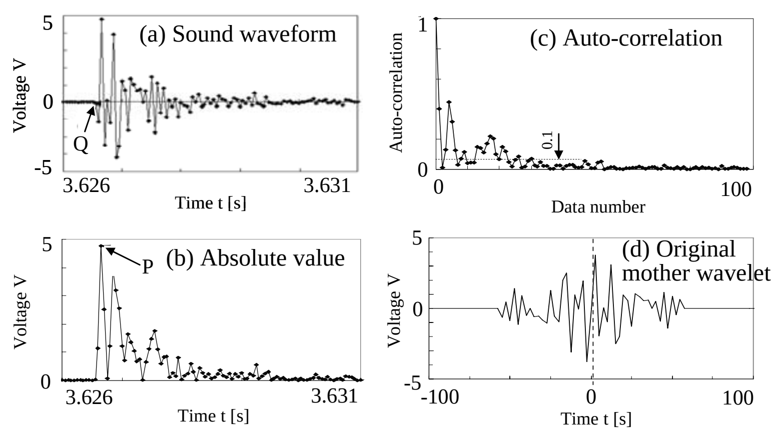


Figure 6 New approach of designing the mother function from the original waveform

3.4 Experimental Verification of Wavelet Analysis

3.4.1 Composition of test wall and experimental method

A test concrete wall where several tile deterioration patterns as shown in Figure 4 had been reproduced was constructed. A schematic picture of the test wall arranged in tile deterioration patterns is shown in Figure 7-(a). The wall face is constructed of mosaic tiles and four patterns (Type①～④) as mentioned in paragraph 3.1 were included into the wall.

Figure 7-(b) shows the finished condition of the tile and inner space. Tile separation was expressed by inserting a non-bonded sheet between the tiles and the bonding layer, and layer separation was reconstructed by making a thin space.

For the experimental apparatuses, an automated hammer device operated by solenoid power was made to strike the tile face once exactly as shown in Figure 8-(a). The hammer's impact was kept at a suitable intensity by referring to basic experiments. The hammer device could be moved on the slider on the test wall as shown in Figure 8-(b).

In the experimental method of the tile inspection, the hammer strikes the tile face and the impact sound is measured by a small microphone installed near the striking point. The sound data is inputted to a memory recorder through the microphone amplifier and A/D transformation device. The obtained sound waveform is analyzed by the wavelet transformation and the condition of the tile deterioration is diagnosed.

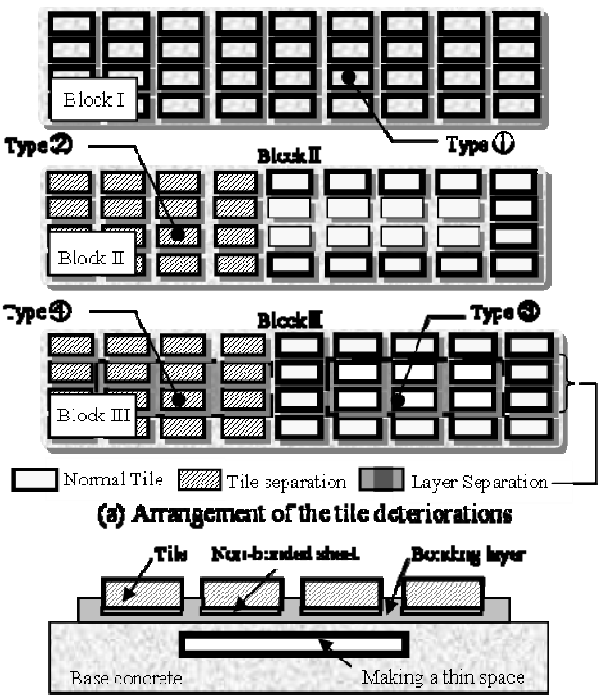


Figure 7 Schematic picture of the test wall arranged the tile deterioration patterns

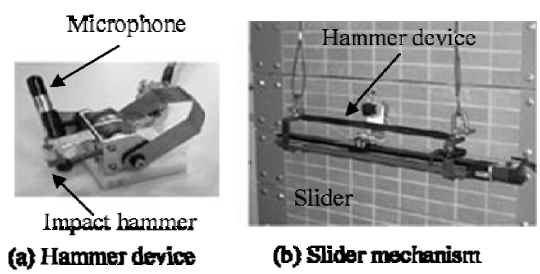


Figure 8 Experimental apparatuses developed hammer device and slider

3.4.2 Visualization of tile patterns of deteriorations by wavelet analysis

Fig.9 shows an example of each tile deterioration pattern visualized on the wavelet plane by wavelet analysis. A sound waveform of the normal tile was used as a mother wavelet function. The following features were confirmed:

- Type ① : The plane of wavelet intensity, was a low level on the whole and the peak wasn't seen in a frequency and time range.
- Type ② : A higher peak appeared in a high frequency range when the hammer was struck. The peak decreased instantly and the plate became almost flat.
- Type ③ : A moderate rise appeared in the high frequency range when the hammer was struck. It decreased instantly, but an undulation in a low frequency range continued for a long time.
- Type ④ : As for Type ②, a higher peak appeared in a high frequency range. Further, as for Type ③, it decreased instantly, but an undulation continued in a low frequency range.

Taking these results into consideration, at tile separation, it can be assumed that the peak appears in the high frequency region because part of the tile that separated from the bonded tile oscillated at natural frequency. However, at layer separation, it can be assumed that the undulation in a low frequency range continued for a long time because void resonance was generated in the layer space.

When the tile separation and layer separation took place in the tiles, the appearance of the characteristics on the wavelet plane corresponded to that of the characteristics combining both separations linearly.

Therefore, the existence of each tile separation is independently judged by observing the characteristics on the wavelet plane. In short, the existence of tile separation is detected by the characteristics of the frequency range on the wavelet plane, and that of layer separation is detected by that of the time range. By visualizing the wavelet analysis, the tile deterioration condition could be inferred.

3.5 Quantity Diagnostics with Wavelet Analysis

3.5.1 Inducement of volume function

Figure10 shows an example of a typical wavelet plane displaying the two tile deterioration modes and the related range of their separation. The related range of tile separation is α and that of layer separation is β. To develop a quantitative method for determining the characteristics of ranges α and β, the volume function of wavelet intensity is calculated in the each range as indicated by equation (4) and the quantity diagnostic method is introduced.

$$F = \int_{f_1}^{f_2} \int_{t_1}^{t_2} I(f, t)\, dt\, df \tag{4}$$

Figure11 shows the expected areas of the volume value normalized by the average value of Type ① being a normal tile. The entire areas are distributed over about four patterns of the normal tile area, the tile separation area, the layer separation area, and the combined separation area. From the areas of each tile separation pattern, a horizontal line of the diagnostics value D_α dividing the existence of the tile separation and a vertical line F_β of the diagnostics value D_β dividing the existence of the layer separation is decided. The condition of deterioration can be easily determined by comparing the values of F_α, and F_β.

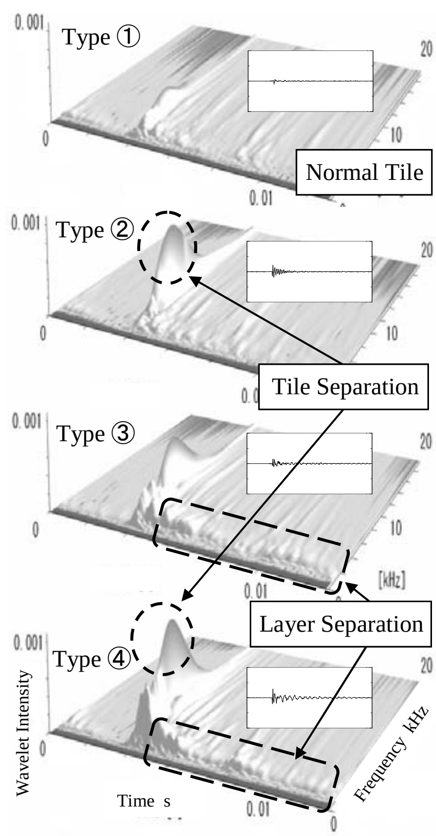


Figure 9 Visualization of each tile deteriorations Patterns on the wavelet plane

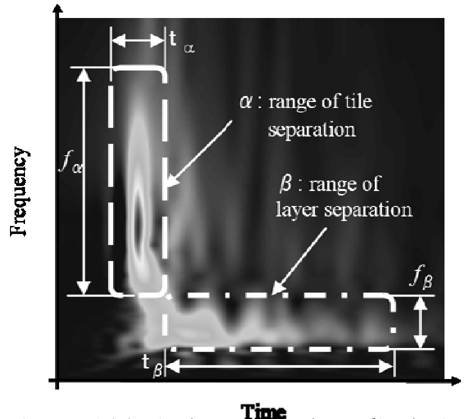


Figure10 Typical wavelet plane displaying two tile separations and related range

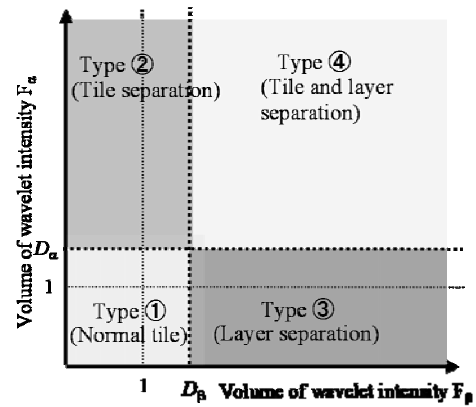


Figure11 Typical wavelet plane displaying two tile separations and related range

3.5.2 Effect of mother wavelet functions and evaluation of quantity diagnostics

In order to evaluate the effect of the mother function for wavelet transformation, the wavelet plane was analyzed by mother functions and the volumes were calculated as indicated before. For the mother function, (a) Gabor-(8), (b) Shannon, (c) Mexican, and (d) Anti-symmetry which were used. (a) Gabor-(8) and (b) Shannon were suitable for determining the characteristics of frequency range, and (c) Mexican, and (d) Anti-symmetry were suitable for determining that of the time range. However, to include the scale effect of the tile itself, for example of mass, length, material etc., the mother function was handmade as mentioned in 3.3.2. (e) Orignal-1 was made by only sound waveform and (f) Orignal-2 was made by processing the natural resonance of the sound.

Fig.12 shows the volume of each wavelet plane corresponding to the choice of the mother functions. The layer separation was able to detect by all the mother functions. However, when using (b) Shannon, (c) Mexican, (d) Anti-symmetry, and (e) Orignal-1, it was difficult to distinguish the tile separation. However, (a) Gabor-(8) and (f) Orignal-2, could diagnose both the tile and layer separation clearly, and especially when selecting (f) Orignal-2, it was understood that the value of compound deteriorations in type ④corresponded to the value that combined both the tile and layer separations linearly. The diagnostic value for the existence of each tile separation was mostly assumed at the D_{α} =1.7 and D_{β} =2.5 quantitatively. A good diagnostic method was established using wavelet analysis and selecting the mother function.

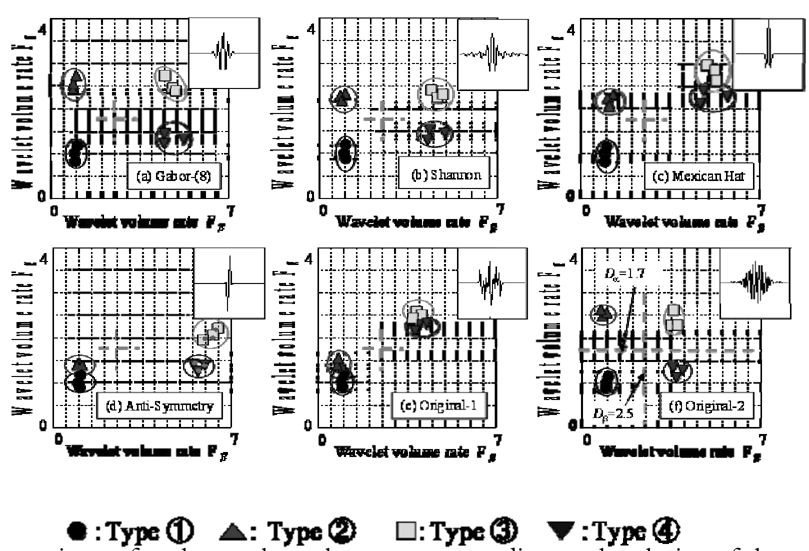


Figure 12 Comparison of each wavelet volume corresponding to the choice of the mother functions and evaluation of the quantity diagnostics

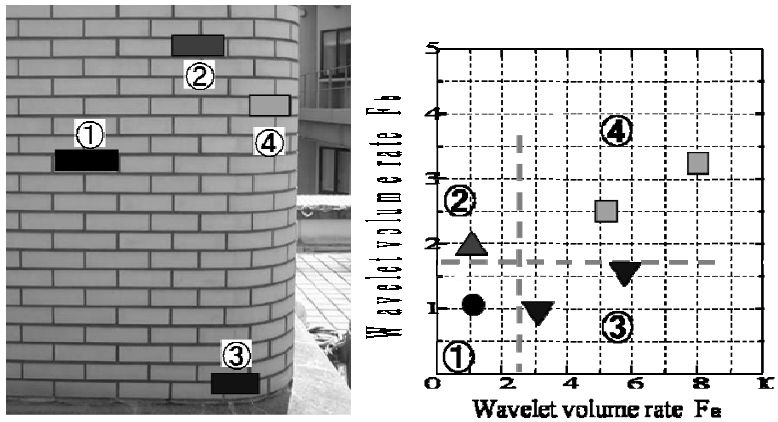


Figure 13 Wavelet volume rate and the position of deterioration by wavelet analysis

3.6 Inspection Result of Actual Tile Wall by Wavelet Analysis

The tile wall of the actual high building was inspected by applying wavelet analysis with the original-2 as the mother function. Figure13 shows the value of the wavelet volume and the position of tile and layer separation. Although the diagnostic of the compound tile deteriorations wasn't detected by the conventional method, each truth was drawn out by the wavelet analysis. Accurately understanding tile deterioration is very useful for the repair technology of the tile wall, as this will suppress accidents related to tile deterioration and improve safety around buildings with tiled walls.

4. Conclusions

An inspection robot and a new diagnostic method using wavelet analysis was reported as a method for inspecting tile deterioration. The following results were obtained.

- The inspection robot can detect a tile's inner condition by a hammering sound. The deterioration result can be instantly determined in real time by two detective parameters. This is a very useful and convenient device for inspecting for tile deterioration.
- In the inspection method for compound deterioration, a new diagnostic method using wavelet analysis was developed. Considering the design of the mother functions and inducing these functions, the most efficient mother function being the original-2 was found by experimental results. Using this function, both tile and layer separation could be clearly diagnosed. A good diagnostic method was established using wavelet analysis.

Finally, the authors thank all who supported the development of the inspection robot and the diagnostics method using wavelet analysis.

References

[1] Inoue, F., and Ohta, Y. (2004). "Research on Automated Judgment of Outer Tile Exfoliation", Proc. of the 10th Symposium on Construction Robotics, pp.335-341.
[2] Inoue, F., and Ohta, Y. (2006). "Research on quantitative Detection of outer tile exfoliation by wavelet analysis",Proc. of the Symposium on the Visualization, pp.34-37.
[3] Chui, C. K. (1992). "Introduction to wavelets", Academic Press New York
[4] Daubechies, I. (1992). "Ten Lecture on Wavelet", SIAM, Philadelphia, 1992.

Construction Automation Process Development – Advancing the Collaboration between Finland and California

Rauno Heikkilä[1], Ty A. Lasky[2] and Kevin Akin[3]

[1]Research Unit of Construction Technology, University of Oulu, P.O. 4400 (Kiilakiventie 1), FIN-90014 Oulu, Finland, rauno.heikkila@oulu.fi
[2]Advanced Highway Maintenance & Construction Technology Research Center, University of California, Davis, CA 95616-5294, U.S.A., talasky@ucdavis.edu
[3] Caltrans Office of Land Surveys, Sacramento, CA 95816, U.S.A., kevin_akin@dot.ca.gov

Abstract

This paper presents the total process model of automation for construction and maintenance, with a focus on roads and bridges. This multi-phase model includes: initial measurements, product modeling and design, construction control and machine guidance, quality assurance and control, and lifecycle operations and maintenance. The paper then provides detailed discussion of current applied research results from Finland and California. The paper gives a summary of the key findings of Finland and California, noting areas of commonality and areas for further investigation. Finally, the paper presents plans for further collaborative research between the University of Oulu and the University of California – Davis.

1. Introduction: The Total Process Model

Automation of infrastructure construction is based on the use of different information models in the different phases of the total working process. For example, the needed machine control models ("control model") can be processed from the related product design ("product model"). The product model is designed and optimized based on measured "initial data models". Constructed structures and products are measured and stored as "as-built models." On-going maintenance and operations of the infrastructure will need different functional measurements ("maintenance model"). In the end-of-product lifecycle, residual value measurements ("residual value model") can be utilized in the long-term planning and development of products and working processes (Heikkilä and Jaakkola 2006). This total process model provides a formal framework for the current and future collaborations between Finland and California.

1.1 Initial Measurements

The automated road construction process starts with on-site input data measurements. In road construction projects, the key input data includes variations in terrain and elevation and soil features. Recently, laser scanning from an airplane, helicopter, or on the ground has developed greatly and become increasingly popular (Hiremagalur *et al.* 2007; Hiremagalur *et al.* 2009). The resulting 3D point cloud adjusted to the relevant coordinate system can be imported into a semi-automatic analysis application that can be used to model not only the contours of the terrain but also tree stands, road alignments, structures, and buildings. However, 3D modeling of underground soil features and conversion into a digitized format is technically far more complicated than above-ground terrain modeling.

1.2 Product Modeling and Design

In product design, a model and instructions for the implementation of the product are created out of the input data. Typically, the products are large 3D objects. If the input data is 3D, design can be carried out 3-dimensionally using CAD tools. The 3D geometric model can be used to produce images for 2D drawings. With improved efficiency in design, it is also essential to be able to make use of the design model for physical construction. If the geometric model is accurate and based on the site coordinate system, it can be used directly for controlling measurements and construction machines. This imposes additional requirements on the quality and accuracy of design.

While geometry is an essential aspect of the model, and represents the focus of the discussion in this paper, other data types must be considered throughout the process. These include road structural layers, pavement characteristics and performance, soil conditions, hydraulic characteristics, existing or anticipated traffic patterns, right-of-way and utility issues, and operational and maintenance issues and measurements.

1.3 Construction Control and Machine Guidance

In the future, there is a very strong view that machines used in civil construction will be increasingly controlled by automated systems. The most advanced systems presently available permit partial automatic control of the machine blade based on 3D positioning and 3D models. However, the functional performance of the systems varies, usually because the corresponding geometric data is not managed completely. Full control of the 3D geometric data and, in particular, inclusion of other property data in the control system remains as future goals.

Previous research and experiments indicate that the control of construction machines and blades requires active coordination by the operator. In this system, the blade position is adjusted automatically with reference to the control model, permitting the operator to select among various options to optimize the process according to the situation at hand. Evidently, pre-calculated paths of travel can seldom be followed.

Typically, 3D point, curve and triangulated surface model models provide sufficient geometric control data for finishing surfaces. However, their data content is not sufficient for a control model for work operations such as the reinforcement of the road bed or stabilization of structural layers, where the objective is to modify the properties of the materials.

1.4 Quality Assurance and Control (QA/QC)

QA/QC is an essential part of the construction automation process. Modern sensing technologies, including 3D laser scanning, robotic total stations, and a variety of pavement quality sensors, can be combined with the CAD data and models used for machine guidance (Kilpeläinen, Heikkilä and Parkkila 2007). Real-time monitoring as well as post-production validation will be necessary to assure the appropriate levels of quality, and compliance with job requirements and specifications.

1.5 Lifecycle Operations and Maintenance

For the safest and most efficient operations and maintenance of the infrastructure, data- and model-driven methods must be applied throughout the lifecycle. Numerous advanced sensing, software, and robotics and automation technologies have been developed and proven by the AHMCT Research Center (AHMCT Research Center 2008). For example, mountain pass road opening benefits from sensing and driver assistive displays based on a GPS/GIS model of the roadway (Yen *et al.* 2008). Several of these systems are now commercially available and are being deployed into infrastructure operations and maintenance (AHMCT 2008; TRAF-tech 2008).

2. Finland Applied Research

Finland is quite active in the field of Road Construction Automation, with on-going cooperative efforts between the University of Oulu, Tekes, the Finnish Road Administration, and numerous private partners. A sampling of these efforts is provided here.

2.1 In the field of Road Construction the results of the Development of an Overall Functional Process Utilizing 3D Data Models and Automation for Rehabilitation of Road Structures – CASE VT4, Haurukylä–Haaransilta (3D-ROAD) project implemented in 2006–2009 in Finland

The main objective of the 3D-ROAD research project is to develop a 3D functional process for enhancing road structures, utilizing automation, and improving the efficiency and quality of the measuring, design and construction processes (see Figures 1-3). The partial objectives enabling achievement of the main objective are: to complete the research tasks required for the development of an overall functional process; to model the overall functional process in order to complete a full-scale test project; and to complete an experimental implementation of automation, i.e. the overall functional process, in an actual construction project.

The Finnish Road Administration (FRA) was also participating in the research project in co-operation with the University of Oulu, aiming to develop subphase-specific requirements to speed up the implementation of new technology as part of the 3D-ROAD project. The requirements were be included in the requests for bids, enabling the actors to present their own implementation solutions as part of their bids. The test project gave the companies an opportunity to develop and test automation technology.

The common objective of the University of Oulu and the FRA was to complete the construction on the actual test site, setting a global standard in utilization of progressive 3D automation. The project was completed using the best technology and skills available today. The first aim was to get the companies to start developing their automation technology. The second aim was to acquire more comprehensive and accurate initial data utilizing accurate 3D positioning. The third aim was to design and model new 3D road geometry as part of the structural rehabilitation project. The fourth aim was to utilize 3D machine control in the construction of the pilot project.

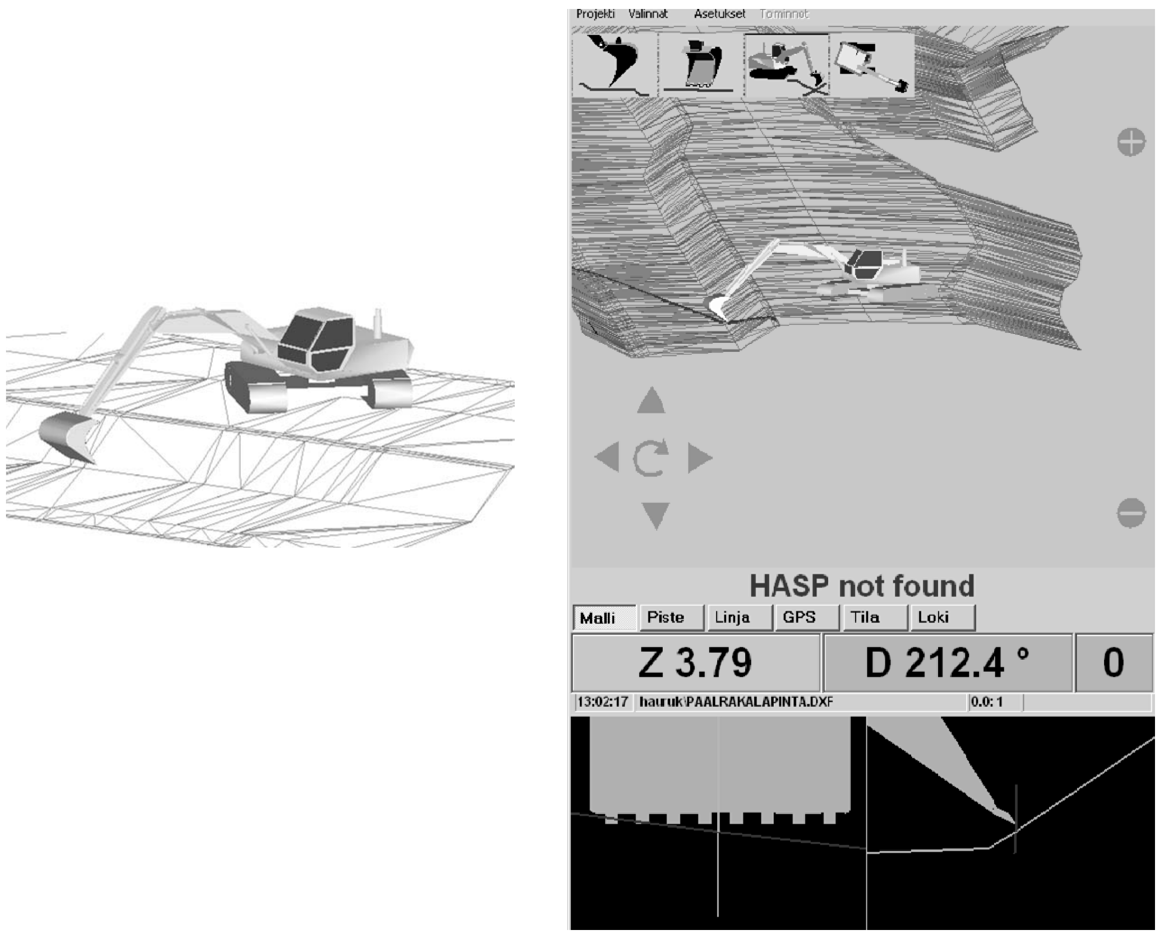


Figure 1: Examining the quality of machine control models in the user-interface of 3D measuring application software

Finland has about 150 3D machine control systems currently in use. Software used includes Tekla Xstreet, Novapoint, and Bentley Inroads for road design tasks. Vehicle-based laser scanning systems have been tested and used, and helicopter-based laser scanning has been in wide use. Finally, 3D GPR measuring has been tested and partly used in construction projects.

2.2 In the field of Bridge Engineering the results of the Bridge Product Modeling and Construction Automation Development (5D-SILTA) project implemented in 2004–2007 in Finland

5D-SILTA was an umbrella project under which various separate R&D projects were carried out in cooperation between the actors in the consortium (Pulkkinen, Karjalainen and Heikkilä 2008). In the project, 5D technology refers to the production, transfer and utilization of data in more than three dimensions throughout the total operating process of bridge building. The research and development focused on the development of 3D laser scanning and GPR scanning of bridges, the transfer of measurement results and road geometry to the 3D product modeling of bridges (see Figure 4), the development and improvement in the efficiency of road design based on product modeling, the diverse utilization of product model data in quantity surveying and cost accounting, scheduling, and procurement planning and management of construction work, and in 3D measurements used to control and validate (QA/QC) construction.

Figure 2: 3D control of excavators in the test of 3D ROAD

Figure 3: 3D control of excavators in the test of 3D ROAD – user interface

Three-dimensional laser scanning enriches the geometric measurements taken during various stages of bridge construction and improves their efficiency and precision. By utilizing the developing technique of bridge scanning, the measurements of the starting data can be expanded further and thus produce more information for the planning of bridge repairs. In bridge design, the shift to 3D product modeling increases the efficiency and speed of design work, reduces design flaws, improves change management, and directly serves the various aims of visualization. A contractor can utilize a product model directly in quantity surveying and bid calculations, procurement operations, schedule management, and even in measuring. Thus, 5D technology integrates the designer and contractor into a more interactive and continuous process of cooperation. The client can utilize the product model when checking plans, for example, and in later maintenance and repair stages.

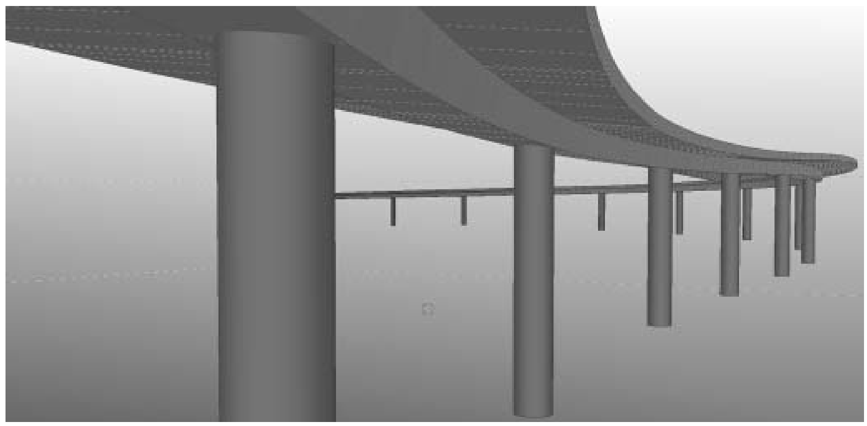


Figure 4: An arched concrete slab modeled in a Finnish test (Tekla Structures)

3. California Applied Research

The California Department of Transportation (Caltrans) current design process is to produce 2D drawings for roadway and bridge construction by in-house engineers. A small portion of the designs are done by engineering firms. Roadway designs are developed using a combination of CAiCE (Autodesk, San Rafael, CA.) and Bentley MicroStation (Exton, PA.) Bridge designs are done using in-house modified and commercial software. The completed 2D plans are used for bidding and building the projects. Contracts are awarded through a competitive bid process and contractors are by law not allowed to be preselected. This prevents the contractor from collaborating with the designer before the project is awarded. Although 3D design data exists, it is often not in a complete and edited 3D model. The creation of a full 3D model is not a required design product at this time.

Caltrans began receiving requests for electronic design data, in addition to the 2D plan sheets, from contractors in 2001-2003. Vendors also demonstrated the use of their equipment for construction automation. However no data was presented, then or now, to conclusively show that automated construction saved time and/or money on highway projects. Construction automation was seen by some Caltrans employees as just another way to use the data that was currently being produced and no workflow change was needed.

One project that attempted to use construction automation was on I-15 near Barstow (see Figures 5 and 6). New lanes were being added to the existing roadway in both directions for 15 miles. The paving subcontractor attempted to use construction automation to control the concrete paver. Two total stations were tracking retro-reflective prisms on the paver and providing XYZ positioning. The weather was hot (air temperature 105^{o} F) and the concrete was setting up very quickly, the paving crew was inexperienced with the equipment, and the contractor wanted both grade stakes and control for the total stations. This additional survey control created demands on the survey department that they could not meet. Ultimately, the contractor returned to using the traditional grade stakes on the project. This combination of environment and experience factors limited the effectiveness of the equipment and showed that the technology might not work in all cases.

In 2003 a Machine Guidance Committee was formed by members of the Caltrans Divisions of Design, Construction, Computer-Aided Design and Drafting (CADD), Office Engineers, Legal, and Surveys. The committee discussed construction automation and it became apparent that there were several challenges.

Caltrans current road design software does not use a full 3D model during design. The software relies on cross-sections as the basic design tool. Full 3D lines are developed through a separate process once the design is finished. Lines are not interactive with the cross-sections, and if a flaw is discovered they must be recreated. New road design software is being evaluated but has not yet been purchased. A change in software should make it easier to create files needed by the contractors.

Figure 5: Concrete paver, automated construction on I-15 near Barstow, CA

There were many concerns about delivering electronic data to contractors. The electronic files might not match the 2D plans, additional claims could be based on the discrepancies, the lines created from cross-sections may contain defects, unknown translation errors could be introduced by different software, and no design file will ever be perfect. These concerns have been addressed but still remain a concern to engineers. Non-standard Special Provisions have been added to contracts to allow delivery of electronic files to contractors but this has been on a case-by-case basis.

The committee developed a set of guidelines based on delivery of electronic files after award (Caltrans 2005). A pilot project was developed in District 11 on the second stage of a bypass around the town of Brawley in Imperial County (Caltrans 2004). State Routes 78 and 111 are being rerouted to reduce traffic on city streets. Originally the electronic design files were to be provided to bidders on the project before award. Due to concerns by the Division of Construction the files were delivered after award. Construction did not start until June 2007 and is expected to be completed by 2010. A “lessons learned” document will be completed when the majority of the work is done.

Figure 6: Robotic total stations tracking retro-reflective prisms on the paver and providing XYZ positioning

Another pilot project was started in 2008 in District 3, around the small town of Tudor, in Sutter County on State Route 70. The design files were included in the bid package and the contract was awarded. Major construction has yet to begin.

There have been many projects done with machine automation around California. Contractors have been given existing electronic design files, or the created their own from existing 2D information. These projects have been poorly documented because of the time pressure to get the projects built. Caltrans management does support developing policy for the use of machine automation.

Caltrans is also looking to other DOTs, the American Association of State Highway and Transportation Officials (AASHTO), and the Transportation Research Board (TRB) to share information. Caltrans has an active representative on AASHTO's Automated Machine Guidance (AMG) Technology Implementation Group (AASHTO 2008). A Request for Proposal to identify guidelines for the use of AMG was sent out and was due on 12/2/2008 (NCHRP 2008). The finished study should be valuable to all DOTs. The TRB subcommittee AFB 80 (Geospatial Data Acquisition Technologies in Design and Construction) is also investigating machine automation (TRB 2008).

In conclusion, Caltrans is investigating machine automation, but the real champions of this new technology are our construction contractors. The major challenges have been producing an electronic design file, changing existing workflows, developing new contract language, purchasing better road design software, overcoming organizational and personal resistance to change, and documenting the results. A compelling case for cost and time savings to Caltrans would make all the difference.

4. Conclusions and Future Research

The total process model for automation for construction and maintenance is expected to provide substantial benefits for transportation infrastructure. Research, development, and pilot testing in both Finland and California have demonstrated the feasibility and many of these benefits, and have also identified areas for further investigation and improvement. The model will allow us to identify key features and requirements for Caltrans and its contractors to appropriately apply machine guidance and construction automation. As noted with respect to California's results, a critical need is a detailed business case analysis demonstrating the cost-benefit for machine guidance and the use of a total process model for the complete infrastructure lifecycle.

The AHMCT Research Center at the University of California and the Research Unit of Construction Technology at the University of Oulu have identified a number of areas for cooperative research in construction automation, machine guidance, and model-driven lifecycle operations and maintenance of transportation infrastructure. We plan to pursue a carefully orchestrated pilot project with Caltrans, structured in a manner to properly evaluate and document the benefits and costs of these technologies in a well-controlled environment. In addition, we are investigating application of GPR for the pro-active evaluation of bridge deck health. We also intend to collaborate on the application of 5D technologies, and the evaluation of their benefits to the DOT. Finally, we plan to develop an on-going series of workshops bringing together government agencies, researchers, and industry, to accelerate the application of innovative and beneficial automation technologies for highway construction, operations, and maintenance.

Acknowledgments

The authors thank their research partners for their assistance in the efforts related to this paper. Rauno Heikkilä thanks the Finnish Road Administration and Tekes, along with many partner companies, especially Tekla Corporation, Novatron, WSP Finland, Destia, and Professor Timo Aho for his great support and contributions in this long-term research. Ty Lasky and Kevin Akin thank the Caltrans Office of Land Surveys and the Division of Research and Innovation for their on-going support of automation efforts, as well as Professor Bahram Ravani, Professor Steven A. Velinsky, Mr. Kin Yen, and the entire AHMCT Research Center team, for their contributions and background information in automation lifecycle research.

References

[1] AASHTO (2008), "AASHTO Technology Implementation Group - Automated Machine Guidance," http://tig.transportation.org/?siteid=57&pageid=689

[2] AHMCT (2008), "Automated Roadway Debris Vacuum System," http://ahmct.ucdavis.edu/index.php?title=DebrisVacuum
[3] AHMCT Research Center (2008), "The Advanced Highway Maintenance & Construction Technology Research Center," http://ahmct.ucdavis.edu
[4] Caltrans (2004), "SR-78/SR-111 Expressway (Brawley Bypass)," http://www.dot.ca.gov/dist11/facts/78-111.pdf
[5] Caltrans (2005), "Interim Guidelines for Machine Guidance Technology," http://www.dot.ca.gov/hq/row/landsurveys/
[6] Heikkilä, R. and M. Jaakkola (2006). "Automation of Road Construction - the State of the Art in Europe," International Symposium on Automation and Robotics in Construction, Tokyo, Japan.
[7] Hiremagalur, J., K. S. Yen, K. Akin, T. Bui, T. A. Lasky and B. Ravani (2007), "Creating Standards and Specifications for the Use of Laser Scanning in Caltrans Projects," AHMCT, Davis, Final Report UCD-ARR-07-06-30-01, June 30.
[8] Hiremagalur, J., K. S. Yen, T. A. Lasky and R. Bahram (2009). "Testing and Performance Evaluation of Fixed Terrestrial 3D Laser Scanning Systems for Highway Applications," Transportation Research Board 88th Annual Meeting, Washington, DC.
[9] Kilpeläinen, P., R. Heikkilä and T. Parkkila (2007). "Automation and Wireless Communication Technologies in Road Rehabilitation," International Symposium on Automation and Robotics in Construction, Kochi, Kerala, India.
[10] NCHRP (2008), "RFP 10-77: Use of Automated Machine Guidance (AMG) within the Transportation Industry," http://www.trb.org/TRBNet/ProjectDisplay.asp?ProjectID=2504
[11] Pulkkinen, P., A. Karjalainen and R. Heikkilä (2008). "5D Bridge – Development of 5D Process and Tools for Bridge Engineering," IABSE (International Association for Bridge and Structural Engineering): Information and Communication Technology for Bridges, Buildings and Construction Practice, Helsinki, Finland.
[12] TRAF-tech (2008), "Automatic Cone Distribution for Safety," http://traftech.com/products.php
[13] TRB (2008), "Annual Meeting, Session 404: Machine Control: Hurdles for Contractors and Departments of Transportation," http://www.trb.org/am/ip/assembly_detail.asp?id=14080&e=445&pre_ff=
[14] Yen, K. S., M. T. Darter, T. A. Lasky and B. Ravani (2008). "Development and Field-Testing of a GPS-Based Mountain Pass Road Opening Driver Assistance System," Institute of Navigation GNSS Conference, Savannah, GA.

Optimal Manipulation Trajectory and Laying Pattern Generation Algorithm for Handling Robot

S. H. Lee[1], S. N. Yu[2], S. J. Lim[2], B. G. Ryu[2] and C. S. Han[3]

[1]Master course, CIM & Robotics Lab., Department of Mechanical Engineering, Hanyang University, Seoul, Rep of Korea; PH (82)31-400-4062; FAX (82)31-406-6398; e-mail: hopezic@gmail.com
[2]Ph.D. Candidate, CIM & Robotics Lab., e-mail: hymecer@gmail.com
[3]Professor, CIM & Robotics Lab., Division of Mechanical and Information Management Engineering, Hanyang University, Ansan, Rep of Korea; PH (82)31-400-4062; FAX (82)31-406-6398; e-mail: cshan@hanyang.ac.kr

Abstract

The conventional brick handling such as road paving site, building construction site is allocated to the construction workers. Since these kinds of tasks are obviously labor-intensive and tedious, there have been many approaches to the automation at the construction site. However, the automation of the block-laying task still has several problems caused by the poor surroundings - inequable construction materials and uneven working conditions like various laying and paving patterns. Herein, this paper proposes an integrated mobile manipulator system operated by the optimal laying pattern and trajectory generation algorithm. The pattern generator is designed by the “Fast Algorithm” based on Steudel’s algorithm; the trajectory generation algorithm is based on the “Overlap Method” which is a treatment skill for robot-surrounded obstacles. This study mainly shows the efficiencies of the proposed pattern and trajectory generation algorithm for the brick-laying task and the performance evaluation of the prototype system.

Keywords: Brick Laying Pattern Algorithm, Manipulation Trajectory Generation Algorithm, Brick-laying, Brick-paving, Mobile Robot

1. Introduction

Automation in the industrial field, in specific, factory automation has made good progress. Operators used to be included in a conventional manufacturing process with uniform working conditions - a formal production line. Automation outside the production line, however, has several limitations and difficulties in adapting to actual conditions, manual or semi-automatic machining tools are mostly used in modern industries. Automated machines or robots are starting to work with man on the jobsite, specifically in a construction site. Through many reports, brick laying and paving is the laborious task with harsh repetitive working conditions; they often cause fatal injuries. Miedema and Vink (1996a, 1996b) found that the highest workload is experienced by the bricklayer when the bricks are located 0~50cm above the work floor and the highest workload of the bricklayers’ assistant is seen in loading and unloading process. [1] To solve these problems, Anliker (1988) developed one of the earliest prototypes of the semi-automated bricklaying masonry machine which is able to build pre-assembled brick walls up to 8 meters long. [2] And Pritschow (1996) proposed a brick laying robot which can operate such a picking bricks or blocks task on the construction site from prepared pallets, applying bonding material, and erecting brickwork with high accuracy and quality. [3] However, these studies commonly have several critical drawbacks. First, they did not consider the importance of an optimized brick-laying pattern generation; hence, the constructor should design the laying pattern of the wall or load separately and check the possibility of the robot to perform the laying task (Fig. 1). Second, they did not pay attention to the motion optimization of the robot arm based on the brick laying position and surrounding obstacles. Motion and trajectory optimization can increase the efficiency of the entire task. This paper defined the brick-laying task and a couple of assumptions for the system configurations. Next, the main pattern generation algorithm - the “Fast Algorithm” is introduced briefly. Then, the designed trajectory generation algorithm to travel between the gripping point (initial point) of picked and palletized point (target laying position) is explained; its performance is verified.

Fig. 1 Conventional method and proposed robotic automation method

2. Concept Design of the Proposed System

2.1 Definition of Target Task

Brick-laying task firstly spreads mortar on sides or ends of brick already laid. The end of the laid brick is buttered with mortar and shoved against the preceding brick, spreading mortar on both sides of closure brick. Laying the brick into position and squeezing the mortar to a width of 10mm, the surplus mortar is extruded from the bed. Then, it is scraped off with the trowel; the same process is repeated onto the next. Since this repetitive characteristic accompanies harmful effects to the workers, the goal of the research was set up for assist these task.

2.2 Configurations of the Proposed System

Generally, the performance of the robotic system and the efficiency of the associated task are determined by adequately sharing the entire task between human and robot, and setting the limitation of the role of the robot system. The aim of this study is to design an assistive robot system which will help construction workers with the most repetitive, and monotonous procedures of brick laying, but not substitute them. Due to many kinds of uncertain field conditions, this study specifically limits the role of the robot to carrying the piled blocks and delivering to the target position by a generated optimal pattern and trajectory. As shown in Fig. 2, in the brick-laying procedure, squeezing the mortar, tapping down the brick, and extruding the surplus mortar from the bed are carried out by the construction workers; it shows the predefined working procedure of each task using the brick laying robot while considering the design strategy. In this study, the unit pattern generation area is defined to use the optimal pattern generation algorithm in the brick- laying robot. The maximized pattern generation unit area for every single stop, considering the robot motion range, is shown in the ($L \times W$) area of Fig. 2. The junction point is the area which the robot cannot cover using its own motion range. Therefore, the bricks belonging to that area are laid by the human worker simultaneously, while the robot performs the brick-laying task. If the user can define the laying timing sequence of various types of bricks, the complex laying is possible. Fig. 3 shows the data process of the entire system. In the following chapter, the Fast algorithm for the brick-laying pattern generation as an initial point of the OLP simulator and real robot system is briefly introduced.

3. Optimal Pattern Generation Algorithm

3.1 Definition of the Fast Algorithm

The "Fast algorithm" gets similar processes with Steudel's algorithm in generating the initial four solution patterns [4]. In addition, Treatment 3 is adopted to apply the heuristic to the central hole in the following three methods, recursively, so as to remove the overlapped area (Fig. 4).

(1) The 1st method: the bricks are cut at the two horizontal edges of the overlapped area.

(2) The 2nd method: the bricks are cut at the two vertical edges.

(3) The 3rd method: the bricks are cut at the left vertical edge and the lower horizontal edge.

As there are not any considerations about all of the block sizes in this algorithm, computing time should be ensured in the whole process. The initial solutions of the first phase find the combination, and define 4 parameters as follows. There are three fundamental rules for the second phase, that I just for

(1) $\bar{a}$ *: When maximizing the length of the block and disposing the bricks lengthwise, the maximal possible number of bricks* = $5l$.

(2) $\underline{a}$ *: When maximizing the length of the block and disposing the bricks lengthwise, the minimal possible number of bricks* = $2l$.

(3) $\bar{b}$ *: When maximizing the width of the block and disposing the bricks lengthwise, the maximal possible number of bricks* = $8w$.

(4) $\underline{b}$ *: When maximizing the width of the block and disposing the bricks lengthwise, the minimal possible number of bricks* = $2w$.

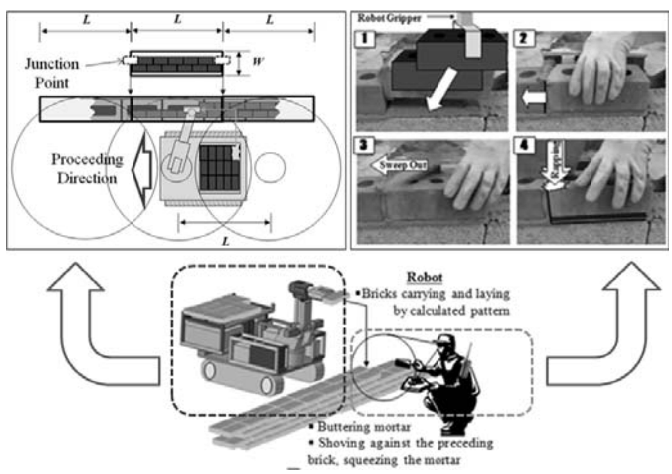


Fig.2 Brick-laying task with proposed robot system: Robot and Human

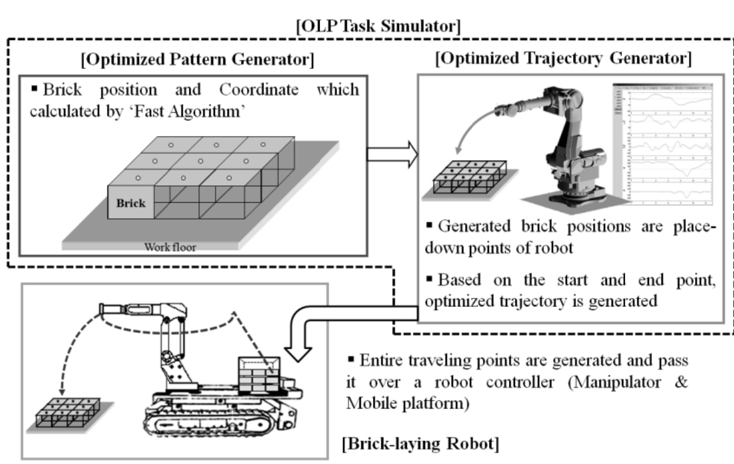


Fig. 3 Proposed system based on OLP task simulator

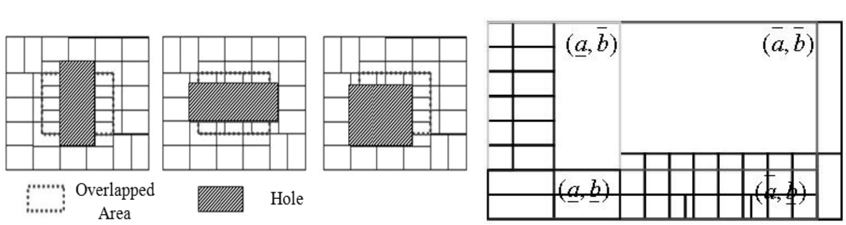


Fig.4 Treatment of the Fast algorithm and its parameter

In the first phase, (L_1, W_1), such as $(\bar{a},\bar{b})$, $(\bar{a},\underline{b})$, $(\underline{a},\bar{b})$, and $(\underline{a},\underline{b})$ are combined, and (L_1, W_1), the width and length of the other blocks can be determined.

$$(L_2,W_2)=(L_4,W_4)=\left(\left[(L-L_1)/w\right]w,\left[(W-W_1)/l\right]l\right) \quad (1)$$

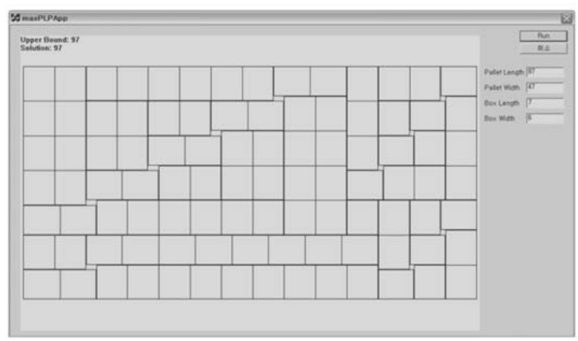
$$(L_3,W_3)=(L_1,W_1). \quad (2)$$

3.2 Computational Performance of the Fast Algorithm

The proposed algorithm is implemented on Visual C++ 6.0 and compiled with maximized-speed option. This algorithm test generated a 2D pattern of bricks and its calculation speed. As a hypothesis, the load balancing of the brick and its stability were not considered. The results of Table 2 were acquired by a computer with a K6-350-

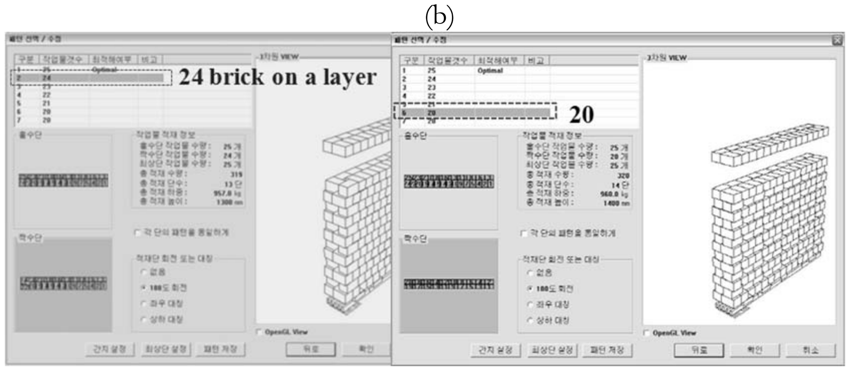
(L, W, l, w)	Number of loaded bricks
(1000,1000,205,159)	30
(1000,1000,200,150)	33
(22,16,5,3)	23
(30,22,7,4)	23
(14,10,3,2)	23
(53,51,9,7)	42
(34,23,5,4)	38
(57,44,12,5)	41
(87,47,7,6)	97
(1200,800,176,135)	38

L	Length of laying plane
W	Width of laying plane
l	Length of brick
W	Width of brick



(a) Result of Fast Algorithm and pattern generation simulation in 2D case

(b)



(b) Brick-laying pattern generation simulation in 3D case

Fig.5 Developing the prototype brick-laying pattern generation simulator

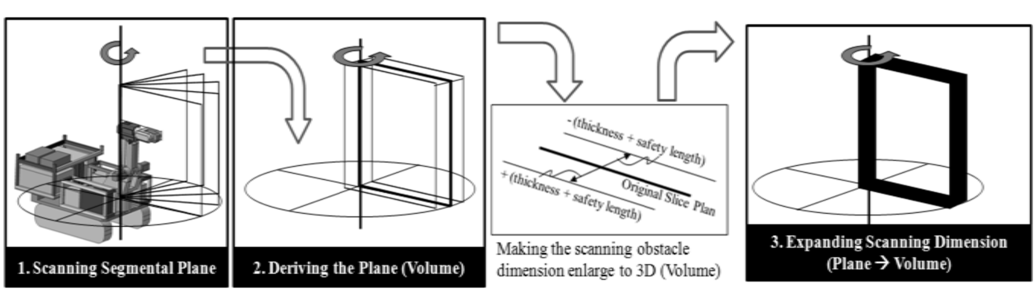


Fig.6 Expanding dimension of the segmental scanning plane

To use this algorithm practically, one dimension of height is applied additionally, and the 3D brick laying simulator is realized on like the Fig. 5 above.

4. Synthetic Trajectory Generation Algorithm

4.1 Configuration Space (C-Space) and A Algorithm: Mapping Obstacles on C-Space*

Couple of main components of the system such as palletized bricks, battery packs, and main controllers can be obstacles at the same time for the brick-laying system. Hence, the concept of C-space to solve this systematic obstacle problem was applied. The configurations defined unknown variables that exactly express the position and direction of an object, and the C-space represented all of the spaces [6]. The coordinate for each configuration was defined; each point in this coordinate that was approached to the robot gripper was expressed by the robot joint angles (configuration, posture) of the proposed system. Fig. 10 shows an example of the generation of configuration space. First, on the basis of the joint of the base frame, the imaginary plane was rotated 360 degrees. In this movement, the objects surrounding the robot were scanned, and an outline of the section was generated. The left side of Fig. 6 describes the specified brick-laying robot layout. The outline, including its interior, could be considered an obstacle. In this study, the outline was acquired by using an end effecter of the robot, and the free-movement and obstacle zones in the C-space were generated as shown in Fig. 6. To help distinguish the 3D shape of the C-space, various brightness intensities and colors are used. This figure is necessary to generate the optimal path using the A* algorithm described in the next paragraph.

4.2 Consideration of the Real Size of the Robot for Trajectory Generation

4.2.1 Modified Slice Plane

One of the disadvantages of the A* algorithm is the required computing time. The aforementioned approach considers the robot arm as a bar. Hence, the computing time load is relatively low. However, a real industrial robot has an original volume, and these factors have to be applied to the A* algorithm. The next step was to consider the real volume of the robot when it scans obstacles and generates C-obstacles. To do this, the slice planes were redefined because it was assumed that the original slice plane had no thickness, the modified slice plane had a thickness, and the factor that changed the scanning point of an obstacle of each angle was a group of both sides of the boundary of the modified slice plane (Fig. 6). The thickness of the plane was determined by the real size of the robot arm, including its gripper and load.

4.2.2 Graham's Algorithm for Convex List

Fig. 7 shows the scanning points that used the modified slice plane. The proposed system used factors of convex list point of objects and the sum of half the thickness and a safe distance. Convex list was generated by the inside apexes of objects and intersection points. If the number of intersection points was less than two, the slice plane is regarded as meeting with an apex or edge. Graham's algorithm was used to generate the convex hull which was used as the new boundary of the object when the modified slice plane was applied [7].

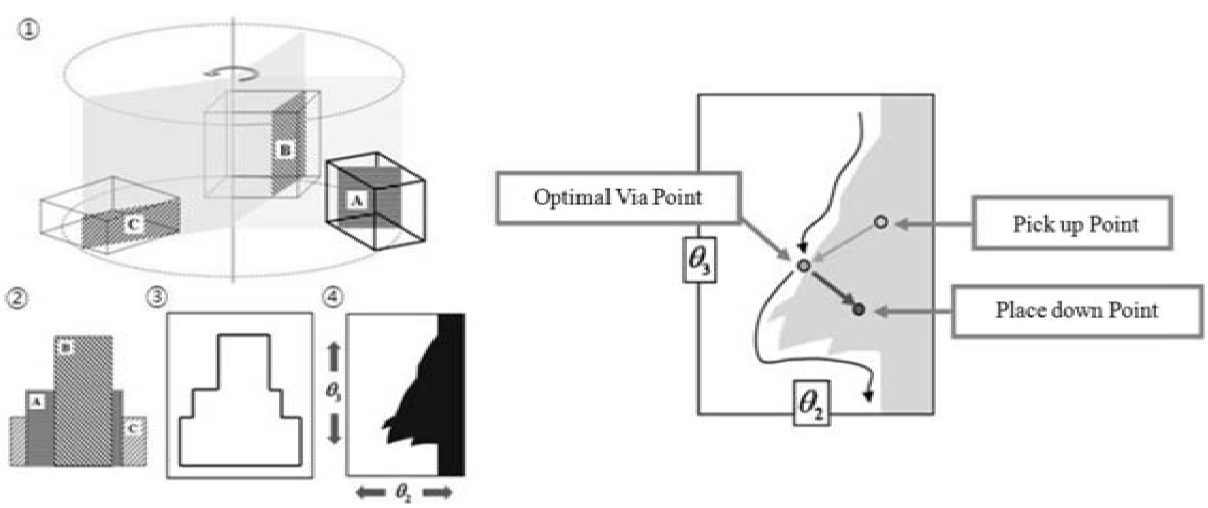


Fig.7 Procedure of Overlap Method and trajectory generation with θ_2 and θ_3

4.3 The Overlap Method for the Brick-laying Trajectory Generation

4.3.1 Basic Concept

The computing load is a critical problem in the area of software development. The main purpose of this study, as described in the Introduction, is to develop an OLP (offline programming) simulator specific to brick-laying automation. If the real size of a brick-laying robot is considered to generate the optimized trajectory, an A* algorithm is a relatively expensive method. To use this algorithm, the C-space has to be generated, but this requires a large amount of computing load. To focus on the characteristics of the brick-laying task, the new strategy which is devoted to the generation of the set of boundaries (convex) of the obstacles was proposed. As shown in Fig. 13, the proposed method overlaps the scanned images of each brick at one plane and obtains the outer line of the overlapped image. This method used the total traveling distance from the pick-up point of the bricks to the place-down point via the outer line of the overlapped area. The following equation was used to optimize the traveling distance that the robot must negotiate to deliver one brick from the pick-up to the place-down point:

$$T_{opt} = A[abs\{(P_{via} - P_{pick-up})_{\theta_2}\} + abs\{(P_{place-down} - P_{via})_{\theta_2}\} + B[abs\{(P_{via} - P_{pick-up})_{\theta_3}\} + abs\{(P_{place-down} - P_{via})_{\theta_3}\}] \quad (3)$$

4.3.2 Considerations of the Via Point

The robot path, however, is not composed of only three points (a place-down point, an optimal via point, and a place-down point). Therefore, this algorithm is exhausted to find an extra via point that would travel the whole path, from the start to the end point. To do this, the optimal via point is used as the initial point. If the gripper of the robot reaches this point, a collision between the gripper and the obstacle can be avoided by changing θ_1. The definition of the collision or gap between the robot and place-down point and the obstacle is decided beforehand.

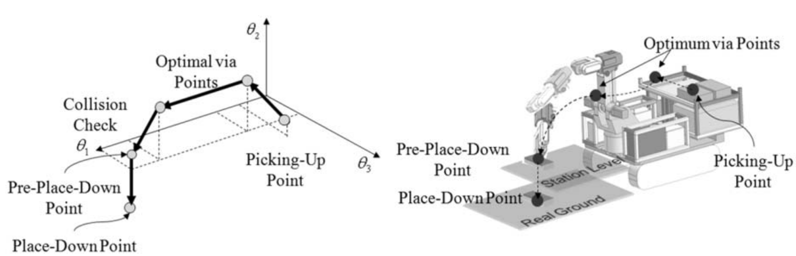


Fig.8 θ_1 for optimal via point generation

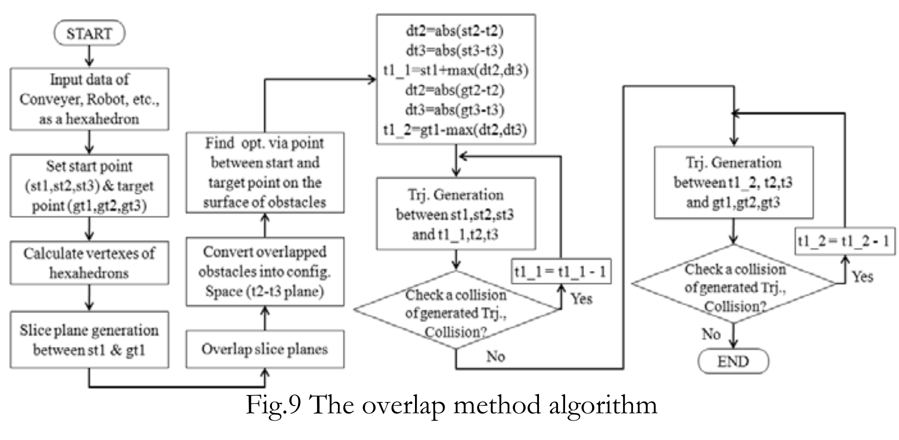


Fig.9 The overlap method algorithm

The place-down point is calculated based on the base frame plane of the manipulator. Therefore, when the robot performs the brick-paving task, the place-down point has to be considered first. Through these several treatments, the intermediate via points are decided as shown in Fig. 8.

(1) st1, st2, st3: θ_1, θ_2, θ_3 *of the starting point*

(2) gt1, gt2, gt3: θ_1, θ_2, θ_3 *of a goal point*

(3) t2, t3: θ_2, θ_3 *of an optimal path point*

(4) t1_(i): θ_1 *of an optimal path point (ith iteration)*

The total travel points are composed of [picking-up point] → 1st optimal via point, [via(θ_{11}, θ_{12}, θ_{13})] → [▪▪▪] → nth via point, [via(θ_{n1}, θ_{n2}, θ_{n3})] → final optimal via point, [via(θ_{f1}, θ_{f2}, θ_{f3})] → [place-down point]. Here, i and j of θ_{ij} mean the ith generated via point of the jth joint of the robot manipulator. Fig. 9 shows the detailed algorithm of the overlap method. This one deals with every surrounding obstacle in every unit step of the process ("unit step" means one cycle of pick-and-place task). As the shapes of the obstacles, especially palletized bricks, are changed at every step, this approach takes strong advantages on that is easy to calculate the path.

5. Result and Conclusions

To prove the efficiency of the proposed methodology, all types of trajectory generation methods described in this paper are simulated, and the results are compared. As shown in the graph, the computing time of A* algorithm that considered the volume of the robot is remarkably different depending on the situation encountered at every step (Fig. 10). However, the overlap method produces fast and stable computation results regardless of the place-down position and configurations of surrounding obstacles. Fig. 10 shows the prototype brick- laying system and simple performance tests. This is about brick positioning performance, and each position is calculated by the proposed OLP simulator based on the "Fast Algorithm" and "Overlap Method." In the next study, there will be trials for synchronizing the mobile platform with the operating algorithm to perform the brick-laying and paving task while continuously moving. Then, there will be trials for improving the dynamic performance of the mobile platform such as minimizing the hysteresis error, and keeping the constant distance precisely against the wall using by the autonomous navigation system. Lastly, there will be field-tests and make-up for a user-friendly operating interface.

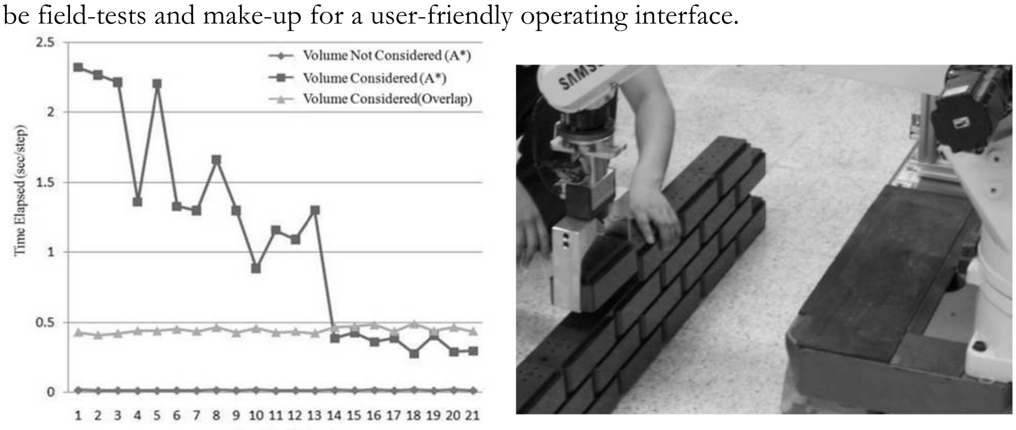


Fig.10 1-step brick-laying task and brick positioning demonstration (Brick, 21EA)

Acknowledgement

This work was supported by R&D Program of MKE(Ministry of Knowledge Economy), Research Fund of Hanyang University (HYU-2008-T) and a grant from Construction Technology Innovation Program(CTIP) funded by Ministry of Land, Transportation and Maritime Affairs (MLTM),SRC/ERC program of MOST (grant # R11-2005-056-03003-0).

References

[1] M. Miedema and P. Vink, "Metselen met de rug rechtop: ergonomische verbeteringen in de bouw [Bricklaying with the back straight: ergonomic improvements in the construction industry]," Tijdschrift voor Ergonomie, pp.38-44, "Fluitend opperen [Whistling during mechanical transport of bricks]", Uitvoering, Conference of National Society of Ergonomics, pp.154-162, 1996.
[2] Anliker, F.J. (1988). "Needs for Robots and Advanced Machines at Construction Sites", ISARC(1988), Tokyo, Japan.
[3] G. Pritschow, Dalacker, M., Kurz, J. & Gaenssle, M. (1996). "Technological aspects in the development of a mobile bricklaying robot", Automation in Construction, 5(1), 3-13.
[4] G, Y.G. & Kang, M.K. (2001). A fast algorithm for two-dimensional pallet loading problems of large size. European Journal of Operational Research, 134(1), 193-202.
[5] G. Pritschow, M. Dalacker, J. Kurz, J. Zeiher, A mobile robot for on-site construction of masonry, IEEE/RSJ/GI International Conference on Intelligent Robots and Systems (IROS '94), vol. 3, 1994, pp. 1701–1707.
[6] Warren, C.W. (1993). Fast Path Planning Using Modified A* Method. Paper presented at the IEEE International Conference on Robotics and Automation (ICRA '93), Atlanta, GA, USA
[7] Graham, R.L. (1972). An Efficient Algorithm for Determining the Convex Hull of a Finite Point Set. Info. Proc. Letters, 26, 132-133.

An Overview of Autonomous Loading of Bulk Material

Ahmad Hemami and Ferri Hassani

PhD, Department of Mining and Materials Engineering, McGill University, Montreal, Canada, H3A 2A7, {ahmad.hemai, ferri.hassani}@mcgill.ca

Abstract

Autonomous loading implies a fully automated scenario in which automated excavating machines, such as front-loaders, load themselves from a heap of bulk material and deliver to the dumping site. The process comprises all the functions of loading, navigating, obstacle detection and avoidance and unloading to be automated and controlled by a supervisory computer. Autonomous loading benefits a number of industries such as construction and mining, from economical view point as well as other concerns like operators safety when the workplace is not hazard free. Despite all the benefits and despite considerable amount of research on the subject, there are no commercially available systems that can be purchased and put to work. In addition to a breakdown of all the tasks that need to be automated and the difficulties involved, this paper reviews and reports the various research and/or development activities that have been carried out during the past two decades on the subject.

Keywords: loading, automated, autonomous, excavation, autoloading

Introduction

Over the past three decades a rather significant amount of work has been carried out regarding automation of excavating machinery. It would be fascinating in an earth moving operation to benefit from a similar scenario as available in manufacturing, where robot manipulators are programmed to perform many of the desired tasks. Such a scenario is, however, much more complex than that of manufacturing. Such capability is called autonomous loading, or it can be referred to as autonomous excavation or other similar terms. The complexity stems from the nature of the operation and the fact that there are various excavating machines and that each one must be dealt independently. Moreover, there are a number of tasks involved in autonomous loading and, in order to have a successful operation all of these tasks must be automated.

Figure 1 illustrates a very simple loading site in which there is only one excavator that must transfer the earth from a heap to a single unloading point. The problem would have been much easier if rotary excavators and conveyer belts were employed. But, the problem under investigation is as it is shown in Figure 1, where cyclic excavators such as backhoe or front-end loader have to interact with a medium that is heterogeneous and whose behavior is stochastic.

In the loading scenario shown the following facts must be taken into account and the corresponding problems must be resolved:

1- The heap to load from has a dynamic structure that changes with each cycle of operation. The loading point, thus, is not fixed and needs to be determined for each cycle of loading.
2- The loading must be successful each time no matter what forces the loading tool experiences from the medium. In this case bucket filling is the ultimate objective. In other cases a different objective may be involved, instead. For example when a ditch is dug by a backhoe, in the last run it is necessary to have a flat surface. The safe operation and safety of the machine must always be respected. In addition, the efficiency of the work cannot be overlooked.
3- When the bucket is filled, the excavator must navigate to the unloading point. This implies deciding on the trajectory to follow, while detecting and avoiding any obstacle. The speed of this transport must be reasonable, meaning that it should take about the same (ideally less) time that it takes for an operator.

4- Unloading to the dumping point is to be with no collision with any carrying truck, belt or so on. The whole load must be delivered (In the case of wet soil, for instance, it is possible that a good portion of the material is carried back without being emptied.)
5- Operation (3) must be repeated while a new trajectory is planned and followed, as required.
6- In all the above operations the environment conditions, such as the slope of the terrain and the temperature (and the gravity constant, if not on Earth), play a role in the operation, which cannot be ignored and must be considered.

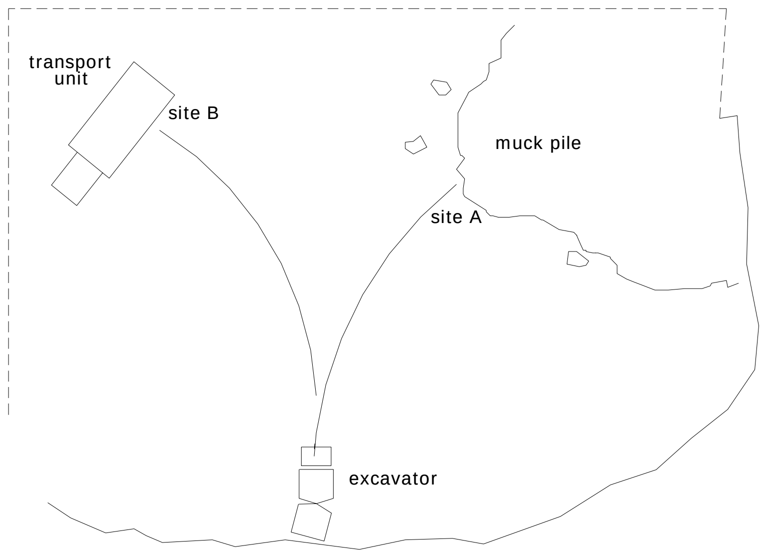


Fig.1- A simple scenario of autonomous loading

A fully autonomous system implies that all the operations are automated and controlled by a supervisory computer. Semi-automated systems, however, are possible and have been practiced for simple cases. For example, if the loading is done by an operator but the navigation between a specific area near the loading site to a specific area near the unloading site is automated. This is useful when there is a long enough distance between the two sites.

Main Tasks

The above example reveals the main tasks involved in autonomous loading. Basically these tasks are:

(a) Excavation (or loading)
(b) Navigation
(c) Obstacle detection
(d) Obstacle avoidance

The main issue in this paper is the first task which, despite numerous works on the subject, has not yet yielded a successful result. Navigation has received tremendous attention, not only as a task for autonomous loading but as part of mobile robots in manufacturing. Mobile robots, contrary to Automated Guided Vehicles (AGV) which have been employed in manufacturing much earlier, have the advantage that they do not required buried wires as well as the fact that they are programmable and, thus, flexible in this sense. Also, the work regarding obstacle detection and avoidance is widely treated in mobile robot technology.

The extension of the mobile robot technology to non-manufacturing application started with the automated navigation of a Load-Haul-Dump unit in an underground mine between a stope (where the fragmented rock is loaded for transportation) and the ore-pass (where the dump-track unloads its load). A roof mounted reflected line and a camera have been used to guide the vehicle along the mine passages (A schematic of a Load-Haul-Dump unit, which from a loading function viewpoint resembles a front-end loader, is shown in the next section). An LHD (Load-Haul-Dump) is a wheeled vehicle which is made merely for underground mines. This work was carried out by Noranda Technology Centre in Canada. This unit was not equipped with obstacle detection and avoidance capability to protect it from the boulders that could have fallen on the route. Later on the reflecting tape was replaced by a laser system.

Excavating machines

Figure 2 shows the most popular excavating machines used in construction sites, road work and mining. Automation of excavation or auto-loading implies that the operation of an excavator is automated and the machine can load its bucket with the proper motion and interaction with the medium. It is obvious that such a capability requires a good deal of artificial intelligence to be embedded into a machine in the form of various sensors, algorithms to interpret the sensory information and control laws based on which the machine moves itself or its bucket. As can be seen, each machine has its own characteristic dictated by its structure, meaning that it must be treated differently from the rest for the purpose of turning to an automated device.

What is common between all the various types of excavating machines is the fact that all must interact with the medium to be loaded or excavated. They have to overcome the resistive force experienced by their bucket when penetrating into a medium. In most of the cases the penetration into the medium is in the form of cutting (and breaking), digging and scooping.

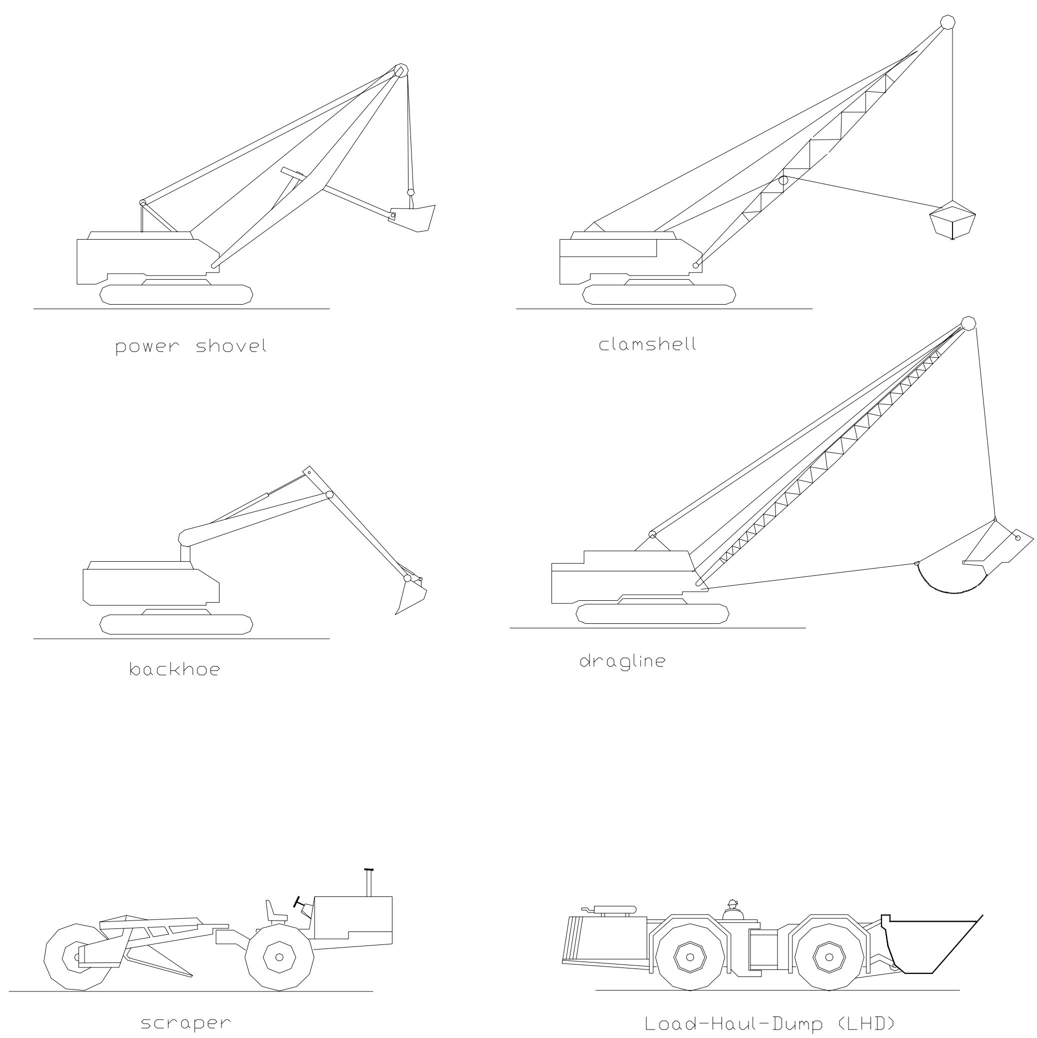


Fig. 2- Various excavating machines

Any automated system works based on the measurement of the main variable(s), forming the error based on the comparison between actual and desired value(s) and a control policy based on which a feedback signal is generated for the control input. The principal question in auto-loading is “What must be measured and what must be controlled". For all excavators, since motion (of the bucket) is involved, a robotic approach (That is, modeling the excavator as a robot manipulator) is an appropriate one. In this sense, the effect of each actuator on the bucket can be defined. If a trajectory for the motion of the bucket is defined, then the variation of each actuator and its speed can be calculated. Furthermore, the relationship between the forces acting on the bucket and those at the actuator level may be defined.

Although it may sound perfect, dealing with an excavator is much more complicated compared to a robot manipulator, because of the following reasons:

1- The forces from the medium to the bucket can have enormous variation not only from one cycle of loading to another, but within the same cycle of operation. These forces can easily become very large at the actuator level and, as a result, any of the actuators may reach its maximum capacity. This is more likely to happen with non-heterogeneous material like a mixture of different soils with different size boulders in it (Fragmented rock, for instance).
2- The cutting tool (bucket) always finds the least resistive path to follow.
3- The structure of the excavator is much more flexible compared to a robot manipulator, specially for the large size payload at the bucket. The body of the machine thus can deform under the influence of large resistive loads mentioned above.
4- As a result of (1) to (3), the trajectory followed by the bucket can significantly deviate from the desired.

In manual operation of any of the excavators shown in Figure 1, and other machines with similar structure, an operator uses his judgment and skills to load the bucket. The motion of the bucket could consist of back and forth maneuvers and repeats. In other words, bucket motion is composed of a series of logically induced smaller motions, and can be very complex. Indeed, an operator judges the forces of medium interaction on the bucket, based on what he sees and hears the load inside or in front of the bucket and the vibrations he feels from the machine. He then moves the bucket accordingly.

Obviously, it is quite difficult if not impossible to enhance a machine with the same type of capability and intelligence as an operator has. Thus a systematic analysis of the way and the extent to which an excavator must be enhanced with sensors and artificial intelligence has been the main focus of the on-going research on the subject of automation of excavation, thus far.

Automated excavation history in brief

The start of the work directly targeted for automation of the process of excavating by one of the non-rotary machines shown in Figure 1 goes back to late 80's and onward, as the reported work starts from early 90's. The prevailing force, one can say, was for both commercial benefits and extraterrestrial discovery aspects [Bernold (1991)]. The earliest research started in Soviet Union as exemplified by the work of Mikhirev (1986). In the west, the work associated with the commercial side are those at Lancaster University, UK [Seward et al (1988), (1992)] and Purdue University [Vähä et al (1991)] towards the automation of a backhoe for digging ditches, and the work at Noranda Technology Centre (NTC), in Montreal, Canada, for automation of both loading and transportation for a Load-Haul-Dump (LHD) unit. LHD is extensively used in underground mines. The nearest category of on-ground excavators to an LHD is a wheeled front-end loader. The work at Lancaster University was more of experimental nature on a small backhoe equipped with force sensors [Bradley & Seward (1995)]. Some success was reported for digging trenches in soft soil. On the other hand, the work at Purdue University involved was of a more analytic nature, such as kinematic and dynamic modeling of backhoe [Cetto and Koivo (1995)]. This analysis is necessary for a robotic approach to tackle the problem. The work at NTC was more involved, since it was directly supported by industry. The approach was to make a machine perform the same sort of motion that a skilled operator did during loading. A real size LHD was equipped with as many sensors that could be accommodated to record the various parameters such as pressure and motions during a loading practice. This was an ad-hoc approach and did not lead to any success. A vibratory motion was later blended into the main motion of the bucket. This led to some success, but it could never be claimed that such a machine could perform as efficient as an operator. The results of the work (or better say, the technical details) on auto-loading were not published at all, as the trend in industry is. NTC was, nevertheless, successful in development of the automated guidance system in underground mines employing a roof mounted reflected tape [Hurteau et al (1992)]. A systematic approach was later begun [Hemami (1993), (1994a), (1994b)] at The Canadian Centre for Automation and Robotics in Mining (CCARM). This work continued with a lower pace later on because of luck of financial support. Another player joining the group later was Carnegie Mellon University [Singh (1991), Luengo (1998), Cannon (1999)] At Robotic Institute a good deal of research on relevant subjects begun [Rowe (1999), Stentz et al (1999a), (1999b)]. . Research activities expanded to University of Arizona [Lever & Wang (1995), Shi et al (1996)] and Tohoku University in Japan [Takahashi et al (1995), (1999)]. More work on the subject includes the research activities in Australia [Corke et al (1997), Roberts et al (1999)].

There are other institutions that are involved in research on this topic, as can be seen from the work of Marshall (2001), Mrad et al (2002), Zweiri et al (2002) Serata et al (2003), Serata et al (2005). The subject still demands further work and funded research. Some progress is continuously reported such as the latest works [Richardson-Little & Damaren (2008), Vahed et al (2008)] though at a slow pace. But strong industry support can speed up the progress.

Relevant research work on excavation automation

In the previous sections a brief but comprehensive introduction to subject was presented. There are a number of topics that are relevant to the subject and the research works by various groups are devoted to one or more of these. It is possible that when an automated system is finally available it will not use all of the findings of the research. The research so far has been on one of the following:

1- Modeling of the excavator as a robot manipulator, or determining the kinematic relationships between the actuator motion and the bucket motion.
2- Determining the force and velocity relationships between the bucket and the actuators.
3- Trajectory studies and planning for the bucket motion.
4- Dynamic modeling of the bucket and actuators.
5- More specific studies of the hydraulics related to bucket motion (Since almost all the excavators work with hydraulic power).
6- Computer modeling of the heap of soil and dynamics of changes due to excavation.
7- Automatic recognition of the shape of the heap and decision making for starting point for loading.
8- Automatic recognition of the environment geometry (Inside a mine, for example)
9- Analysis of the composition of the fragmented rock (size variation) and its effect on loading.
10- Analysis of the forces (various force components) of interaction of bucket and medium.
11- Analysis and formulation of the cutting force and the mechanism of material failure when subject to bucket forces.
12- Use of computer vision to estimate the amount of loaded material.
13- Methods to control the motion (Control strategy).
14- Computer simulation of the process.

The references cited in this paper are examples of the work that has been carried out during the past three decades. By no means can one claim that this review is complete and all the previous works are addressed. From those addressed and their references one can count more than 500 technical papers on the subject. A good review of the various pertinent works before 1997 is presented by Singh (1997).

Since during an excavation task there is an exchange of active and resistive forces, from the bucket and the medium respectively, significant research has been devoted to identifying, analysing and formulation of forces that are involved in the process. Many formulations are proposed, but they have not been verified and are not in accord with each other [Hemami et al (1994c)]. A good review and comparison of the various proposed theories for formulating the earthmoving forces can be found in Blouin et al (2001).

As far as the control philosophy is concerned, there are not many proposed methods to control the process yet, and none has proven to provide an acceptable solution. Indeed, there are only three main approaches. Bullock et al (1990) were the first to propose a two-level control based on the measurement of the excavation forces during the operation. Some researchers believe that a controller does not need to be based on the analysis or measurement of the forces and, therefore, controllers based on fuzzy logic instructions can work [Ha et al (2002), Ha & Rye (2004)]. This second approach was tested earlier by some other groups [Shi et al (1996)]. A third approach is that any action be based on the maximum power of the excavator. In other words, an excavator is always working with maximum output power on the bucket. This is the approach implemented to LUCIE [Seward et al (1992), Bradley & Seward (1995)].

Any proposed method must be proven by implementation to real tasks and in the realistic conditions. In other words, more experimental research is required and any successful approach must also prove to be economically viable.

Summary

This paper deals with autonomous excavation which implies automation of a number of functions related to earthmoving that are currently performed by human operators. It summarizes (1) the necessary functions to be automated in order to benefit from autonomous excavation, (2) the history of the research work that has been carried out towards this purpose, (3) the wide range of research activities that are pertinent to the subject, and (4) a list of references that can be traced for learning the technical content of the related material as well as getting in touch with the people/organizations that are active on the subject.

References

[1] Bernold, L.E. (1991) Experimental Studies on Mechanics of Lunar Excavation, J. of Aerospace Engineering, 4(1), 9-22.

[2] Blouin, S, Hemami, A and M. Lipsette (2001), "Review of Resistive Force Models for Earthmoving Processes", Journal of Aerospace Engineering, Vol. 14(3), 102-111

[3] Bradley, D.A. Seward, D.W. (1995), "Developing real-time autonomous excavation-the LUCIE story" Proc. the 34th IEEE Conf. Decision and Control, Volume: 3, 3028-3033

Bullock, D. M., Apte, S. M. and Oppenheim, I. J., (1990), "Force and Geometry Constraints in Robot Excavation, Proc. Space 90 Conf. Part 2, Albuquerque, 960-969.

[4] Cannon, H. (1999), "Extended Earthmoving with an Autonomous Excavator", MS Thesis, Carnegie Mellon University, CMU Technical Report # CMURI- TR-99-10

[5] Cetto, J.A. and Koivo, A.J. (1995), "Controller for Excavator During Digging Operations", IASTED Int. Conf. Robotics and Manufacturing, Cancun, (Mexico)

[6] Corke, P. I., Winstanley, G. J., Roberts, J. M. (1997) "Dragline modeling and control",

[7] Proc. IEEE Conference on Robotics and Automation

[8] Ha, Q.P., Santos, M., Nguyen, Q, Rye, D. and Durrant-Whyte, H. (2002), "Robotic Excavation in Construction Automation", IEEE Robotics and Automation magazine, Vol. 9(1), 20-28

[9] Ha, Q. P. and Rye, D. C. (2004), "A Control Architecture for Robotic Excavation in Construction", Computer-Aided Civil and Infrastructure Engineering 19, 28–41

[10] Hemami, A. (1993), "Study of Bucket Trajectory in Automatic Scooping with Load-Haul-Dump Loaders", Trans. Institute of Mining and Metallurgy, 102, sec. A, 37-42.

[11] Hemami, A. (1994a), "Study of Forces in the Scooping Operation of a Mechanical Loader", Transactions of the CSME (Canadian Society of Mechanical Engineers), Vol. 18(3), 191-205.

[12] Hemami, A. (1994b), "Modelling, Analysis and Preliminary Studies for Automatic Scooping", Advanced Robotics, Vol 8(5), 511-529.

[13] Hemami, A. Goulet, S. and Aubertin M. (1994c), "Resistance of Particulate Media to Excavation: Application to Bucket Loading", Int. Journal of Surface Mining and Reclamation, Vol 8, 125-129.

[14] Hurteau, R., St-Amant, M., Laperriere, Y., Chevrette, G.(1992) "Optical Guidance System for Underground Mine Vehicles," Proc. International Conference on Robotics and Automation,

[15] Lever P.A. and Wang F.Y. (1995),"Intelligent excavation control system for lunar mining system", J. Aerosp. Eng., Vol. 8(1), 16-24

[16] Luengo, O. Singh, S. Cannon, H. (1998), "Modeling and Identification of Soil-Tool Interaction in Automated Excavation", Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems, Vol. 3, 1900-1906

[17] Mrad, F., Abdul-Malak, M. A., Sadek, S. & Khudr, Z. (2002), "Automated excavation in construction using robotics trajectory and envelop generation", Engineering, Construction and Architectural Management, Vol. 9(4), 325–335

[18] Marshall, J.A. (2001), "Towards Autonomous Excavation of Fragmented Rock:

[19] Experiments, Modelling, Identification and Control", Master's Thesis, Queen's University, Kingston, Ontario, Canada
[20] Mikhirev, P.A. (1986), "Design of Automated Loading Buckets", Soviet Mining Science, 22(4), PP 292-297.
[21] Richardson-Little W. and Damaren C. (2008), "Position Accommodation and Compliance Control for Robotic Excavation", Journal Aerospace Engineering, Vol. 21(1), 27-34
[22] Roberts, J.M., Winstanley, G.J and Corke, P.I. (1999),"Issues for Obstacle Detection for Autonomous Mining and Construction Vehicles", Proc. ISARC 16, Madrid (Spain),
[23] Rowe, P.S. (1999) "Adaptive Motion Planning for Autonomous Mass Excavation", PhD Thesis, Carnegie Mellon University, CMU Technical Report # CMU-RI-TR-99-09
[24] Sarata, S., Osumi, H., Hirai, Y and Matsushima, G. (2003), "Trajectory Arrangement of Bucket Motion of Wheel Loader", Proc. ISARC 2003, Eindhoven (Netherlands), 135-140.
[25] Sarata, S, Weeramhaeng, Y and Tsubouchi, T. (2005), "Planning of Scooping Position and Approach Path for Loading Operation by Wheel Loader", Proc. 22nd International Symposium on Automation and Robotics in Construction (ISARC 2005), Ferrara (Italy)
[26] Seward, D, Bradley, D and Bracewell, R. (1988), "The Development of Research Models for Automatic Excavation", Proc. 5th Int. Symp. on Robotics in Construction, Tokyo, Vol. 2, 703 – 708.
[27] Seward, D., Bradley, D., Mann, J. and Goodwin, M. (1992), "Controlling an Intelligent Excavator for Automated Digging in Difficult Ground", Proc. 9th Int. Symp. on Aut. and Robotics in Construction (ISARC), Tokyo (Japan) 743-750.
[28] Shi, X, Lever, P.J. A and F-Y Wang, (1996), "Experimental Robotic Excavation with Fuzzy Logic and Neural Networks", Proc. IEEE Int. Conf. Robotics and Automation, Vol.1, 957-962.
[29] Singh, S. (1997), "The State of the Art in Automation of Earthmoving", ASCE Journal of Aerospace Engineering, Vol 10(4), 179–188.
[30] Singh, S. (1991), "An Operation Space Approach to Robotic Excavation", Proc. IEEE Int. Symp. Intel. Control., NY, 481-486.
[31] Stentz, A., Bares, J., Singh, S., Rowe, R. (1999a), "A Robotic Excavator for Autonomous Truck Loading," Autonomous Robots, Vol. 7(2), 175-186
[32] Stentz. A. Ollis, M., Scheding, S., Herman, H., Fromme, C., Pedersen, J., Hegadorn, T., McCall, B., Bares, J., Moore, R.(1999b), "Position Measurements for Automated Mining Machinery, Report: Robotics Institute (Carnegie Mellon University), and in Proc. of the 1999 International Conference on Field and Service Robotics
[33] Takahashi, H., Kamata, H., Masuyama, T., Sarata, S. (1995), "Autonomous Shoveling of Rocks by Using Image Vision System on LHD, " Proc. International Symposium on Mine Mechanization and Automation, Golden, Colorado
[34] Takahashi, H., Hasegawa, M. and Nakano, E. (1999), Analysis on the Resistive Forces Acting on the Bucket of a Load-Haul-Dump Machine and a Wheel Loader in the Scooping Task., Advanced Robotics, Vol 13(2), 97-114.
[35] Vähä, P.K., Skibniewski, M.J. and Koivo, A.J. (1991), "Kinematics and Trajectory planning for Robotic Excavation", Proc. Construction Congress 91, Cambridge (MA), 787-793.
[36] Vahed, S.M., Song, S., Dai, J. S., Lam, H. K., Seneviratne L.D. and Althoefer K. (2008), "The Shortest Path of an Excavator Bucket in its Working Zone", Proc.17th IFAC World Congress (IFAC 08), Seoul (Korea), 13822-26
[37] Zweiri, Y.H., Seneviratne, L.D. and Althoefer, K. (2002), "Model-based automation for heavy duty mobile excavator", Proc. IEEE/RSJ International Conf. Intelligent Robots and System, Vol. 3, 2967-2972

Path Re-planning of Cranes Using Real-Time Location System

Cheng Zhang[1], Amin Hammad[2] and Homam AlBahnassi[3]

[1]Graduate Research Assistant, Building, Civil and Environmental Engineering Department, Concordia University, 1455 De Maisonneuve Blvd. West, Montreal, Quebec, Canada H3G 1M8; PH (514) 848-2424, ext. 7074; FAX (514) 848-3171; e-mail: zha_che@encs.concordia.ca
[2]Associate Professor, Concordia Institute for Information Systems Engineering, 1455 De Maisonneuve Blvd. West, Montreal, Quebec, Canada H3G 1M8; PH (514) 848-2424, ext. 5800; FAX (514) 848-3171; e-mail: hammad@ciise.concordia.ca
[3]Graduate Research Assistant, Building, Civil and Environmental Engineering Department, Concordia University, 1455 De Maisonneuve Blvd. West, Montreal, Quebec, Canada H3G 1M8; PH (514) 848-2424, ext. 7074; FAX (514) 848-3171; e-mail: hbahnassi@inframez.com

Abstract

The numbers of reported accidents related to cranes are increasing during the past 10 years. Although simulation and visualization software are available for path planning of cranes, the high dynamic characteristics of on-site conditions often require re-planning the crane's path to ensure safety and efficiency. Any unpredicted objects on site should be detected and tracked in real time and the resulting information should be used for path re-planning. This paper proposes an approach for monitoring and re-planning path of cranes on construction sites using a real-time location system. Data collected from ultra wide band (UWB) sensors attached to the equipment, in addition to an up-to-date 3D model of the construction site, are used to detect any possible collisions or other conflicts related to the operations of the cranes. Re-planning algorithms are discussed to generate a new plan in real time. The advantages of the proposed approach are: more awareness of dynamic construction site conditions, a safer and more efficient work site, and a more reliable decision support based on good communications.

1. Introduction

Safety and productivity issues on construction sites are always among the major concerns of project managers. The complexity of on-site conditions requires careful planning and coordination of different equipment to ensure safety and efficiency. Cranes are one of the most frequently used equipment for lifting objects on site. From 1992 to 2006, there were 323 deaths related to cranes in the U.S. (NIOSH, 2008). These accidents were caused by contact with overhead power lines, struck by booms/jibs, struck by crane load, caught in between, etc. In Canada, there were 56 accidents related to cranes in the province of British Columbia in 2006 (WorkSafeBC, 2008); and during the period of 1974 to 2002, there were 23 accidents with injuries, 26 accidents with death, and 13 accidents with material damage related to cranes in Quebec province (CSST, 2008). Furthermore, the numbers of reported accidents and the resulting deaths are increasing during the past 10 years (Crane Accident Statistics, 2008).

To fulfill tasks efficiently and safely in a complex environment with known and unknown obstacles, ideal conceptual methods are proposed for path planning. During the planning stage, the *model-based approach* is used, where a 3D model of the site is available, which means full information about the geometry of the equipment and the obstacles is given beforehand, so path planning becomes a one-time off-line operation. During the execution stage, the dynamic environment needs another approach, called *sensor-based planning*, with an assumption that some obstacles are unknown, and this is compensated by local on-line (real-time) information coming from sensory feedback (Spong et al., 1992).

Most of the software for crane path planning support only off-line operations, and the main functionality is to locate the crane on site rather than path planning for specific tasks (Cranimation 2007, LiftPlanner 2007). The path planner usually visits the site in advance to check the environment and makes adjustment to the plan. After that, when the crane is carrying on the plan, lifting tasks are usually done through a trial-and-error process, based on feedback provided by the operator's own vision and assessment, hand signals of a designated crane or ground director at the work zone, or radio communication (Arizono et al., 1993). If any obstacle appears, the remainder of the path may need to be re-planned based on the experience of the operator. Few researchers are focusing on real-time crane path planning (Kang and Miranda, 2006). Providing an efficient algorithm for path re-planning can ensure safety and improve efficiency. Figure 1 shows groups of cranes working in parallel in replacing the deck of Jacques Cartier Bridge in Montreal. The crane operator has to cooperate with another crane operator to lift the deck panel together. Since trucks and workers on site may become obstacles on the path of cranes, precisely locating them will significantly reduce the chance of accidents.

Figure 1: Groups of cranes working on a bridge

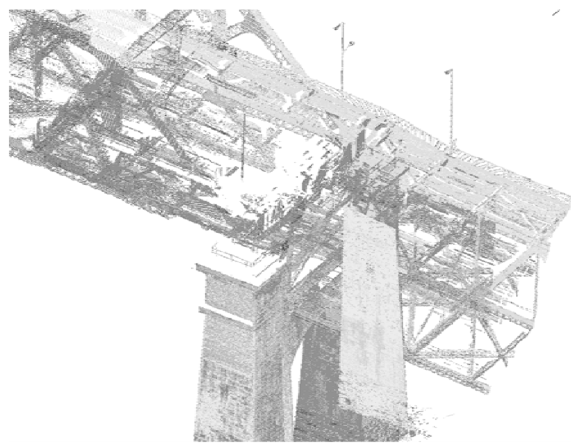


Figure 2: Point cloud collected for a bridge (Mailhot and Busuio, 2006)

This paper is part of a research project at Concordia University concerning real-time planning for construction equipment using multi-agent systems coupled with field data capturing technologies. Details of the overall framework and other enabling technologies can be found in Zhang and Hammad (2007) and Zhang et al. (2008). The present paper focuses on the algorithms for path re-planning and real-time location system (RTLS) using Ultra Wideband (UWB) technology. The objectives of the present paper are: (1) to investigate the usability of UWB tracking technology for cranes; and (2) to test an algorithm for real-time path re-planning based on the criteria of efficient world updates and query updates, and scalability for large number of degrees of freedom.

2. Environment Perception

2.1 Modeling static objects

Several methods are used to create the 3D model for static objects. Photogrammetry is used for calculating geometric properties of objects based on photographic images (Photogrammetry, 2008). Geographic Information Systems (GIS) based 3D modeling is also used to create an urban model based on extruding polygons representing building footprints in maps according to the heights of the buildings (GIS for Archaeology, 2008). These data are becoming more available in some cities. However, these models include mainly buildings and miss other small objects, such as traffic signs, fire hydrants, and electric poles and lines. Therefore, researchers are trying different technologies to create an accurate 3D model of the construction site. 3D laser scanners are used to collect point clouds, which can be transformed by software tools into volumetric objects, representing a precise 3D model including all the buildings and other objects. Repeated work should be carried out to update the model in real-time. Gordon and Akinci (2005) collected data using a 3D laser scanner to support inspection and quality control on construction sites. Figure 2 shows a sample of point cloud collected for the bridge shown in Figure 1. Also, 3D range cameras have been evaluated for construction applications (Lytle et al. 2005) and some important parameters are indicated to optimize the accuracy and minimize errors (Price et al., 2007). Teizer et al. (2006) have used a 3D range camera to model static and dynamic construction resources.

2.2 Tracking moving objects

Moving objects should be detected and tracked in real time and the resulting information can be used for path re-planning. For instance, the positions of cranes need to be tracked and controlled to avoid collisions. On-board instrumentation (OBI) has been used to collect data about the equipment configuration, or other data which need to be monitored. Navon et al. (2004) have developed a tracking and control system using Global Positioning System (GPS) and OBI to monitor in real-time the activity of major construction equipment. However, GPS is unavailable without direct line of sight from the satellites, and accurate GPS receivers are expensive to install on every moving object on site. Therefore, other tracking technologies have been applied in several research projects, such as infrared, optical, ultrasound, RFID (Radio Frequency Identification), etc. Chae and Yoshida (2008) have discussed collecting data on site using RFID active tags for preventing collision accidents. However, RFID can only give proximity of locations. Recently, RTLS have been applied in various areas, such as logistics and manufacturing. RTLS can track and identify the location of objects in real time using nodes (badges/tags) attached to objects, and devices (readers) that receive the wireless signals from these tags to determine their locations. Figure 3 shows the trend of more precise RTLS using ultra wideband (UWB) technology, which delivers a robust localization with an accuracy of up to 15 cm.

UWB is a wireless technology for transmitting large amounts of digital data over a wide spectrum of frequency bands over a distance up to 230 feet at very low power (less than 0.5 milliwatts). UWB has the ability to carry signals through doors and other obstacles that tend to reflect signals at more limited bandwidths and a higher power. Also, UWB works better with metals than other radio frequency (RF) devices. These advantages make it possible to attach UWB tags on construction equipment and other moving objects on site. In this research, Ubisense products are used to evaluate the usability of UWB technology. Both AOA (Angle Of Arrival) and TDOA (Time Difference Of Arrival) are used to locate tags based on triangulation. As shown in Table 1, the combination of TDOA and AOA provides the highest accuracy by using 2 or more sensors.

Researchers have started to investigate the usability of UWB on construction sites. For example, Teizer et al. (2007) have attached an UWB tag to a crane to track the position of the hook for safety issues. Giretti et al. (2008) have applied experiments on position tracking of workers on site for safety purposes. Construction Metrology and Automation Group (CMAG) is involved in measuring the performance of UWB tracking technology in construction (Saidi and Lytle, 2008).

Figure 3: Trend of a more precise RTLS using UWB (Ubisense, 2008)

Table 1: Combinations of the location method and the results (Ubisense, 2008)

Location method	**Number of sensors detecting tag**	**Other information required**	**Result**
Single-sensor AOA	1	Known height of tag	2D horizontal position (+ known height)
AOA	2 or more	None	3D position
TDOA+AOA	2 or more	None	3D position (highest accuracy)
TDOA only	4 or more	None	3D position

3. Path Re-planning Algorithms

There are a large number of algorithms available for generating collision-free paths. Varghese and his colleges have been studying crane path planning for a long time. They have tried different algorithms, such as A*, and Genetic Algorithms (GA), for optimizing the path of cooperative lift with two cranes (Sivakumar et al., 2003). In the research of Ali et al. (2005), a GA algorithm is used and compared with the A* algorithm and the former is considered as a better solution for two cranes working together. However, they assumed that the site contains only static obstructions, and the proposed solutions only provide off-line planning, rather than real-time control.

Finding the lowest-cost path through a graph is central to many problems, including planning for construction equipment (e.g. cranes). If arc costs change during the traverse, then the remainder of the path may need to be re-planned. This is the case for sensor-equipped crane with imperfect information about its environment. As the equipment acquires additional information via its sensors, it can revise its plan to reduce the total cost of the traverse.

During re-planning, the crane must either wait for the new path to be computed or move in an uncertain direction; therefore, rapid re-planning is essential. An efficient re-planning algorithm should be able to plan optimal traverses in real-time by incrementally repairing paths to the crane's state as new information is discovered. Re-planning algorithms focus on the repairs to significantly reduce the total time required for the initial path calculation and subsequent re-planning operations. Re-planning algorithms generalize path planning for dynamic environments, where arc costs can change during the traverse of the solution path. (Stentz, 1995).

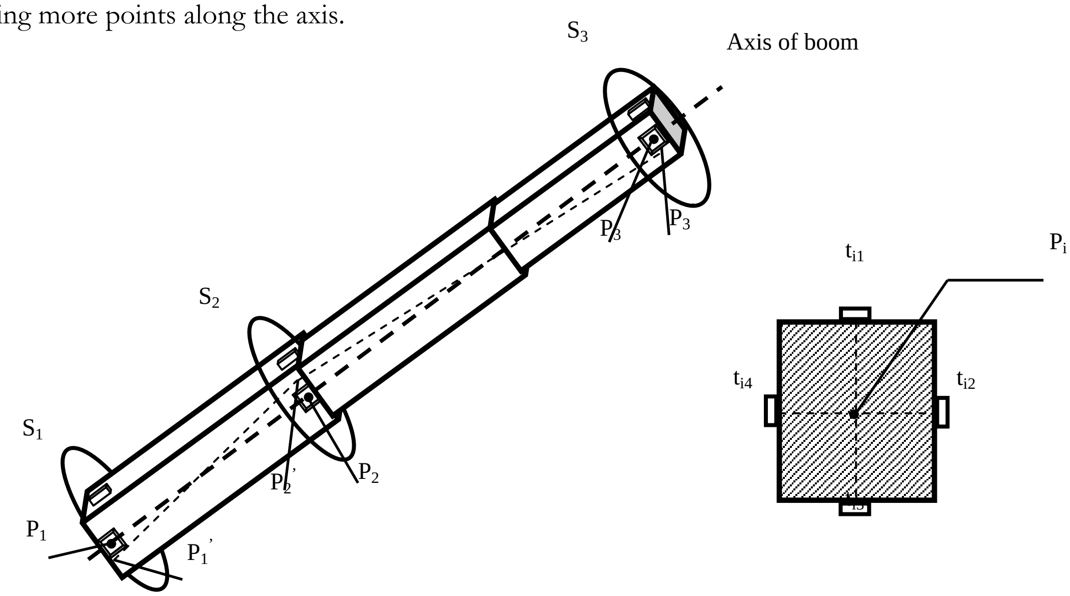
Deterministic re-planning algorithms, such as D*, efficiently repair previous planning solutions when changes occur in the environment. They do this by determining which parts of the solution are still valid and which parts need to be recomputed. However, as the dimension of the search space increases, for example, in the case of multiple cranes working together, deterministic algorithms such as D* simply cannot cope with the size of the corresponding state space. On the other hand, randomized approaches such as Rapidly-exploring Random Trees (RRTs) are a good choice for solving this problem since they are not crippled by its high dimensionality. RRTs have been shown to be effective for solving single-shot path planning problems in complex configuration spaces by combining random sampling of the configuration space with biased sampling around the goal configuration. RRTs efficiently provide solutions to problems involving vast, high-dimensional configuration spaces that would be intractable using deterministic approaches. Brandt (2006) has made a comparison between A* and an RRT algorithm for motion planning of robots, and found that RRT is much faster than A*.

However, in cases where the initial information available concerning the environment is incomplete or the environment itself is dynamic, typically the current RRT is abandoned and a new RRT is grown from scratch. This can be a very time-consuming operation, particularly if the planning problem is complex. For this case, researchers have started implementing the Dynamic RRT (DRRT) as a probabilistic analog to D* for navigation in unknown or dynamic environments (Ferguson et al., 2006). DRRT depends on repairing the current RRT when new information concerning the configuration space is received instead of abandoning the current RRT entirely. It efficiently removes just the invalid parts and grows the remaining tree until a new solution is found.

4. Proposed Approach

During the planning stage, 3D environment model is assumed to be available using methods explained in section 2.1. Based on the updated environment model, a collision-free path plan for a crane is generated using DRRT. During the actual construction work, multiple UWB tags are attached to different components of cranes and other moving objects and workers on site to monitor their position and orientation.

For tracking the location of a hydraulic crane, enough tags should be attached to its different components. Tags can be attached to the first section and the tip of the boom for easy installation and to avoid damaging the tags. Furthermore, tags should be located to fulfill the visibility, orientation, and accuracy requirements. Figure 4 shows a schematic boom with three sets of tags (S_1, S_2, S_3) attached to it. Each set includes four tags (t_{i1}, t_{i2}, t_{i3}, t_{i4}) fixed on each face of the boom, which ensures visibility by the UWB sensors when the boom rotates. The location of the center point of a cross section (P_i) can be calculated based on all or some of the four tags' locations of tags S_i. The orientation and the length of the boom can be obtained by connecting two axis points, P_1 and P_3. The purpose of adding the second sets of tags (S_2) is to get a third point (P_2) on the axis of the boom and to reduce the error of the axis location, and increase the accuracy by having more points along the axis.



(a) Schematic representation of tags on a boom (b) Cross section of the boom with set of tags (Si)
Figure 4: Locations of tags on the boom

The sensors should be set in a way to utilize their antenna pattern which is +/- 90 degrees in the azimuth and +/- 50 degrees in the elevation. The maximum range of sensors can be potentially up to 200 ft. Considering that the boom of the crane will be raised up during work, the sensors should be facing up to capture the movement of the boom. However, other movements, such as the movements of trucks and workers carrying tags attached to their hardhats, may be out of the vertical range of these sensors. Therefore, another set of sensors facing down is needed to capture different objects. Calibration of the sensors should be done to obtain an exact location of sensors with respect to a local coordinate system. Using a reference tag with a known location, the pitch and yaw angles of each sensor can be calculated. The roll angle is set to 0 because the sensors are levelled during installation.

After attaching tags to objects, information about the configuration of equipment and other moving objects are collected and used for collision prediction. If obstacles are detected, the re-planning algorithm is triggered, and a revised path is generated if necessary to guide the movement of the equipment using the following steps: (1) Send signals to the crane for which the path is blocked to stop the movement; (2) Differentiate the type of the obstacle, e.g., equipment or worker, to decide the priority of movement; (3) Check whether the obstacle is for short or long duration (e.g. shorter or longer than 5 min.) based on the goal and the plan of the moving objects; (4) If it is a short duration obstacle, inform the crane to wait until the obstacle moves, then resume executing the plan; (5) If it is a long duration obstacle, then re-plan the path of the crane to avoid the obstacle; and (6) Execute the new plan.

5. Case Study

For the implementation part, we started developing a prototype of the system integrated with DRRT algorithm in "Autodesk Softimage" 3D software package with its 3D visualization and animation capabilities. Motion Strategy Library (2003), which includes variations of RRT algorithms, is used as a base library for developing an integrated motion planning solution in Softimage. The DRRT algorithm is implemented based on the RRT-Connect algorithm (Kuffner and LaValle, 2000).

A laboratory test was applied to a scaled model of a tower crane to simulate a lifting task with a tag attached to the hook. Raw data are collected using the UWB system software showing the tag's name, date, time, and the x, y, z coordinates. These data are transferred to the Softimage through the BVH motion file format, which indicates the hierarchy of the equipment and the motion at each time step. A continuous collision detection is applied to check whether a potential collision exists. Re-planning is triggered when an obstacle appears on the path and a new path is generated based on DRRT algorithm. Figure 5 shows a screen shot of the visualization environment, where the traces of the movement of the hook is shown. The path is not smooth because of the noise in the collected data, which could be cut using filtering.

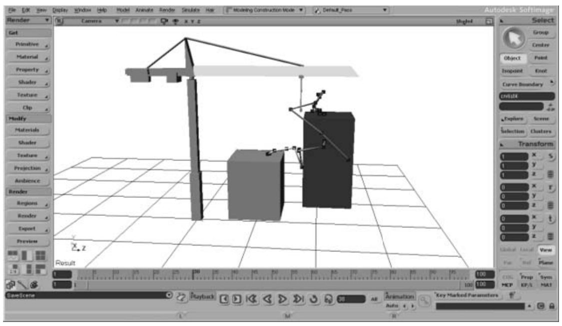


Figure 5: Visualization environment

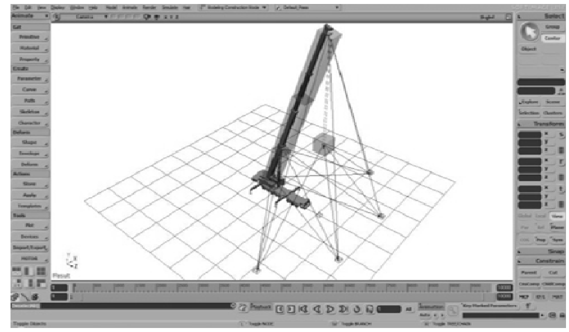


Figure 6: Design of the site setting

An outdoor test was applied to a hydraulic crane which only focuses on the movement of the boom of the crane. Softimage is used to design the setting of the site, as shown in Figure 6. Four sensors are located facing the boom of the crane, and several tags are attached on different parts of the boom. The out-door test was done on the yard of a crane company on December 12, 2008, under - 4°C temperature. Four sensors are installed around the crane to maximize the opportunity that more tags on the boom can get line of sight to at least 2 sensors (Figure 7). The power of the four sensors is supplied by a Power over Ethernet (PoE) switch. To synchronize the signals from different tags, timing cables are used to connect the four sensors. One sensor is acting as a master sensor to receive and synchronize the timing data from the other three. The sensors are calibrated using a tag as a reference point with known position. Tags used in the test are specially designed for use in harsh industrial environments and include several advanced features: an LED for easy identification, a motion detector to instantly activate a stationary tag and a push button to trigger events used by workers. As shown in Figure 8, tags are attached to the boom of the crane at different locations. Two tags are attached to the hook.

Figure 7: Locations of sensors

The result of the test is under analysis and will be reported at the conference. Another out-door test is planned in a warmer weather. Some problems were identified during the preliminary test: (1) Tags are difficult to fix using the adhesive pads on the boom due to the cold weather, and a magnetic mount could be a better solution; (2) Calibration of the sensors is difficult because they were facing up and the reference tag should be at a certain height difficult to reach.

6. Discussions and Future Work

This paper proposed an approach of path re-planning for cranes in real-time using an RTLS. Preliminary test was done on a crane with several tags attached to its boom and hook. DRRT algorithm is proposed for path re-planning and is under implementation. Future work includes: (1) Using multiple sets of sensors to monitor objects with various heights; (2) designing and testing different cases with various sensor locations and tag locations; (3) integrating the path re-planning algorithm and the RTLS.

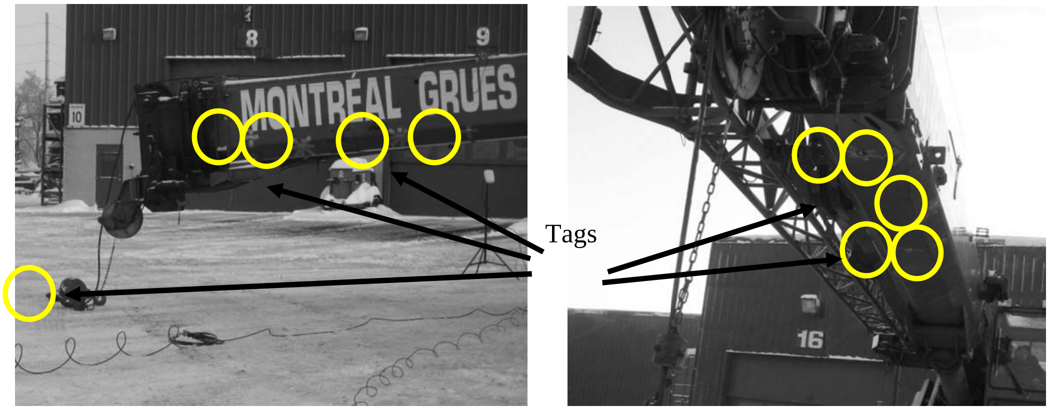


Figure 8: Tags attached to the boom and the hook

Acknowledgement

We would like to thank Mr. Christian Servais from the GUAY Crane Company in Montreal for providing us the opportunity to test the RTLS on a crane; and our colleges for helping in the test setting.

References

[1] Ali M.S.A.D., Babu N.R., and Varghese K. (2005). "Collision free path planning of cooperative crane manipulators using genetic algorithm." Journal of Computing in Civil Engineering, ASCE, Vol. 19, No. 2, 182-193.
[2] Arizono N., Shiota K., Uemura Y., Sasada T., Morikawa Y., and Onoda M. (1993). "Crane operation support system." Res. Dev., Vol. 43, No. 1, 47-50.
[3] Brandt D. (2006). "Comparison of A* and RRT-Connect motion planning techniques for self-reconfiguration planning." Proc., the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, October, China.
[4] Chae, S., and Yoshida, T. (2008). "A study of safety management using working area information on construction site." Proc., the 25th International Symposium on Automation and Robotics in Construction, ISARC, Vilnius, Lithuania, 292-299.
[5] Crane Accident Statistics. (2008). http://www.craneaccidents.com/stats.htm
[6] Cranimation (2006). http://www.cranimax.com
[7] CSST (2008). Commission de la santé et de la sécurité du travail du Québec, http://www.csst.qc.ca/portail/fr/
[8] Ferguson, D., Kalra, N., and Stentz, A. (2006). "Replanning with RRTs." Proc., the IEEE International Conference on Robotics and Automation, 1243-1248.
[9] Giretti, A., Carbonari, A., Naticchia, B., and Grassi, M.D. (2008). "Advanced real-time safety management system for construction sites." Proc., the 25th International Symposium on Automation and Robotics in Construction, ISARC, Vilnius, Lithuania, 300-305.
[10] GIS for Archaeology (2008). www.esri.com/industries/archaeology/business/modeling.html
[11] Gordon C., and Akinci B. (2005). "Technology and process assessment of using LADAR and embedded sensing for construction quality control." Proc., Construction Research Congress 2005: Broadening Perspectives, San Diego, April.
[12] Kang, S. and Miranda, E. (2006).Planning and visualization for automated robotic crane erection processes in construction, Automation in Construction, Vol. 15, 398 -414.
[13] Kuffner, J.J., and LaValle, S.M. (2000). "RRT-Connect: an efficient approach to single-query path planning." Proc. IEEE International Conference on Robotics and Automation, San Francisco, U.S., April.
[14] LiftPlanner (2007). http://www.liftplanner.net.

[15] Lytle A.M., Katz I., and Saidi K.S. (2005). "Performance evaluation of a high-frame rate 3D range sensor for construction applications." Proc., the 22nd International Symposium on Automation and Robotics in Construction, Italy, September, 2005.
[16] Motion Strategy Library (2003). http://msl.cs.uiuc.edu/msl/
[17] Mailhot G., and Busuio S. (2006). "Application of long range 3D laser scanning for remote sensing and monitoring of complex bridge structure." Proc., the 7th International Conference of Short & Medium Span Bridges, August, Montreal, Canada.
[18] Navon R., Goldschmidt E., and Shpatnisky Y. (2004). "A concept proving prototype of automated earthmoving control." Automation in Construction, Vol. 13, No. 2, 225-239.
[19] NIOSH (2008). The National Institute for Occupational Safety and Health, http://www.cdc.gov/NIOSH/
[20] Photogrammetry (2008). http://www.photogrammetry.com/
[21] Price M., Kenney J., Eastman R.D., and Hong T. (2007). "Training and optimization of operating parameters for flash LADAR cameras." Proc., the 2007 IEEE International Conference on Robotics and Automation, Roma, Italy, April, 3408-3413.
[22] Saidi K., and Lytle A. (2008). "NIST Research in Crane Automation: 2007 Overview." the 87th Annual Meeting, Transportation Research Board.
[23] Sivakumar PL., Varghese K., and Babu N.R. (2003). "Automated path planning of cooperative crane lifts using heuristic search." Journal of Computing in Civil Engineering, ASCE, Vol. 17, No. 3, 197-207.
[24] Spong M.W., Lewis F.L., and Abdallah C.T. (1992). Robot Control – Dynamics, Motion Planning, and Analysis, IEEE Press, IEEE Control Systems Society.
[25] Stentz, A. (1995). "The Focused D* Algorithm for Real-Time Replanning" Proc., the International Joint Conference on Artificial Intelligence (IJCAI'95), Montreal, Quebec.
[26] Teizer J., Bosche F., Caldas C.H., and Hass C.T. (2006). "Real-time spatial detection and tracking of resources in a construction environment." Proc., the Joint International Conference on Computing and Decision Making in Civil and Building Engineering, June, Montreal, Canada, 494-502.
[27] Teizer J., Lao D., and Sofer M. (2007). "Rapid automated monitoring of construction site activities using ultra-wideband." Proc., the 24th International Symposium on Automation & Robotics in Construction, ISARC, September, India, 23-28.
[28] UbiSense (2008). http://www.ubisense.net/index.php
[29] WorkSafeBC (2008). http://worksafebc.com.
[30] Zhang C., and Hammad, A. (2007). " Collaborative agent-based system for multiple crane operation." Proc., the 24th International Symposium on Automation & Robotics in Construction, ISARC, September, India.
[31] Zhang C., Hammad, A., and AlBahnassi H. (2008). "Collaborative multi-agent systems for construction equipment based on real-time field data capturing." Next Generation Construction IT: Technology Foresight, Future Studies, Roadmapping, and Scenario Planning (accepted).

Robot-aided Tunnel Inspection and Maintenance System

J. C. González, S. Martínez, A. Jardón and C. Balaguer

Roboticslab Engineering Systems and Automation Department, Universidad Carlos III de Madrid, 28911 Leganés, SPAIN, {jcgvicto, scasa, ajardon, balaguer}@ing.uc3m.es

Abstract

This paper describes an innovative alternative to manual procedures for the application of carbon fiber and resin injection in concrete surfaces in tunnels. It is based on a specially designed light-weight integrated tool for automatic application of Fiber Reinforced Polymer (FRP) and epoxy resin injection. Vision and laser telemeter sensing are integrated into the tool to assure precise inspection and maintenance operations. An interconnection flange allows simple and robust attachment to a robotic manipulator's tip. The robot-tool set is mounted on an articulated lift platform. Therefore, an operator can direct the platform and the robot-tool's operations in a control station placed at ground-level, in a wheeled vehicle on which the articulated lift platform is mounted. An intuitive Human-Machine Interface (HMI) has been developed to allow the operator to identify fissures for the injection of epoxy resin, and to choose where to place the FRP strips. Actual procedures are completely automatic.

Keywords: Robotic automation, Robotic tool design, Tunnel maintenance, Concrete inspection, HMI

1 Introduction

Tunnels nowadays are designed and built to last hundreds of years. However, change in use, new load criterions, and impact and damage caused by natural and human factors can reduce service life to less than a tenth of its original estimate.

Inspection and maintenance operations in tunnels depend heavily upon time and space constraints, the tunnel's intrinsic conditions (reduced space, possible existence of service pipes), traffic flow, and, in railway tunnels, the presence of aerial electric cable. Working conditions in these subterranean infra- structures can be slow and tedious due to dust, humidity, complete absence of natural light, and uncomfortable working conditions. Efficiency and stigma of any operator gradually reduce throughout the day, incrementing the risk index. Any attempt to automate operations performed in subterranean infrastructures will drastically improve short and long-term security, and increase productivity [1].

A great range of factors can cause need for maintenance or reparation operations. Two of the most important of these factors will be treated:

- Fissure formation due to deformation caused by excessive load. Although current legislations tolerate the existence of small fissures, their real dimensions should never exceed a small, reduced range. Reinforced concrete's interior metal infrastructure should never be exposed to ambient atmosphere.
- Loss of quality of the infrastructure's surface due to lack of coating, needed to guarantee correct metal-conglomerate adhesion. Most pathological phenomena (carbonation, chlorine...) pass undetected until external signs appear on the iron structure: rust, section and adhesion reduction...

2 Maintenance Operations

Presently, all inspection and maintenance operations in tunnels are performed manually (Figure 1). Frequently, traffic flow must be cut, and scaffolds mounted, implying the subsequent loss of global productivity. Maintenance operations involve the following set of tasks: superficial preparation, fissure injection, and FRP composite adhesion.

2.1 Superficial Preparation

This includes all of the processes needed to eliminate concrete in bad state. Loss of mechanical capacity or of stability within the rest of the structure can be a possible reason for this need. The surface must be prepared before concrete reparation and/or restoration. Common methods include compressed air blowing, sand abrasion, and hydro-demolition (preparation with high-pressured water).

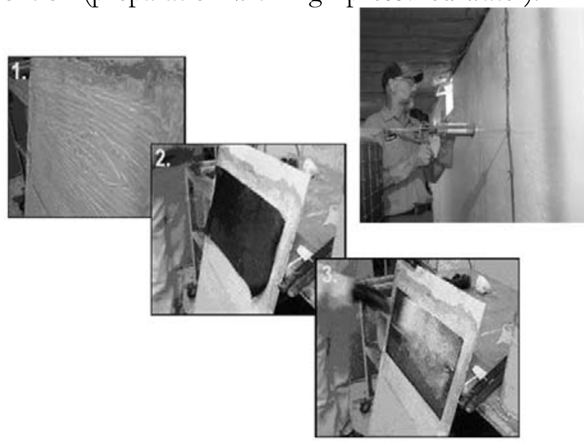


Figure 1. Conventional Manual Procedures

2.2 Fissure Injection

Development of advanced materials such as low-viscosity epoxy resins allows pressure injection of these materials through fissures as small as 0.2 mm thick. They are usually thixotropic, solvent-free, and bi-component. Components are included in different compartments of a same cartridge, and combined through a static mixer at application time. The main objective of these injections is to re-establish continuity in the concrete section. Injections can be used to fill in internal or hard-to-access zones (such as internal nests). This technique has been used and tested on tunnels, bridges, and several nuclear centrals over the past few years.

2.3 FRP Composite Adhesion

Since its first applications in Europe and Japan in the 1980s, superficial reparation and restoration with Fibre Reinforced Polymer (FRP) has progressively increased. Quick and easy installation and high versatility make FRP reparations attractive to owners, engineers and promoters. Carbon or aramid fibres allow high mechanic, thermal, electric and chemical resistance, high modulus of elasticity, and low density [2]. Forms of interest include flexible strips that can be adhered to concrete surfaces using application-specific epoxy resins.

The following considerations should be taken into account when applying FRP to external surfaces:

- Dimension and orientation of each strip.
- Number of strips and sequence of installation.
- Time limits between successive strips.
- Surface temperature and moisture limitations.
- General notes listing design loads and allowable strains in the FRP laminates.

The main strength component of a non-slender, normal-weight concrete member confined with an FRP jacket may be calculated using the FRP confined concrete strength equation (1) or (2).

$$\phi P_n = 0.85\phi\left[0.85\psi_f f'_{cc}\left(A_g - A_{st}\right) + f_y\ A_{st}\right] \quad (1)$$

$$\phi P_n = 0.80\phi\left[0.85\psi_f f'_{cc}\left(A_g - A_{st}\right) + f_y\ A_{st}\right] \quad (2)$$

Confining a concrete surface is accomplished by orienting the fibres transverse to the weakest axis of the member, resulting in an increase in the apparent strength of concrete and in the maximum usable compressive strain [3]. In this orientation, fibres work similarly to conventional spiral or tie reinforcing steel. However, any contribution of the longitudinally aligned fibres to the axial compression strength of a concrete member should be neglected, as fibres are usually displaced in a unidirectional fashion, following the direction of the strip or roll. FRP jackets provide passive confinement, remaining unstressed until dilation and cracking of the member occurs [4]. For this reason, intimate contact between FRP and concrete surface is critical.

FRP adhesion is also used in automotive, marine, and aerospace industries.

Integrated Process Automation

The Robot Aided Tunnel Inspection and Maintenance (RATIM) System is the result of the first attempt ever to automate superficial preparation, fissure injection, and FRP composite adhesion procedures. An integrated system was mounted with a specially designed tool, a robot arm, and a wheeled vehicle with a platform lift (as seen in Figure 2). This configuration has also been applied with Kuka robots in Korea [5], but never before for concrete reparation process automation.

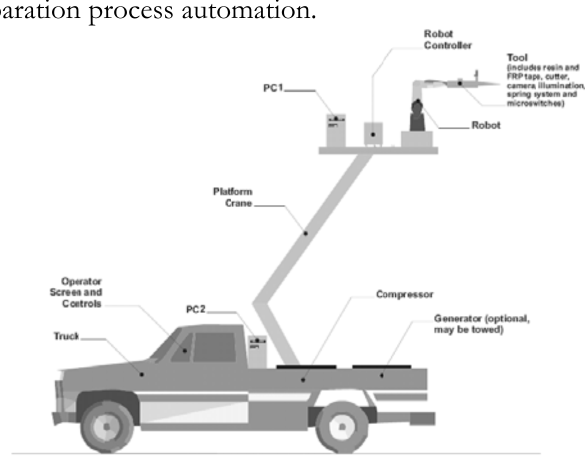


Figure 2. Basic Hardware Architecture

Inspection is performed through a user-friendly guided HMI, where camera image stream is displayed and operation procedures are solicited. Visual servoing based on the depth measure captured by the tool's laser telemeter, and operation-oriented actuators are coordinated through task-specific control software [6].

For precise operation, the robot arm chosen for the application was the Mitsubishi PA-10, a 7 DOF manipulator with a 10 kg load capacity and a 1 meter maximum extension range. The global increment of this range is achieved by mounting the Robot-Tool a 5 meter extensible articulated lift platform. The HMI is installed in the wheeled vehicle's cabinet to which the articulated lift platform is attached. Power for the system can be supplied from an on-board generator, the wheeled vehicle's motor, or the tunnel's basic provided services.

When the surface to treat is inside the reach volume, user extends the vehicle's stabilizers, directs the platform towards the concrete surface, and then follows the HMI's on-screen instructions.

Specially Designed Tool

A light-weight sensed integrated tool was designed and manufactured for the automation of the described processes. Material provider's workshops and real construction sites were visited to analyze manual procedures 'in situ'. Conclusions were that radical and innovative approaches would have to be taken to lead this system to complete automation.

- Minimum changes should be needed to be effectuated on the tool to change between superficial preparation and resin injection process, and superficial preparation and FRP adhesion process.
- A three-in-one, one-pass FRP adhesion process (resin, FRP, resin) is desirable.
- Maintaining the weight of the tool low throughout the design process would allow it to be attached and used within a wide range of commercial and non-commercial robots.
- The robot-tool interconnection flange must allow simple mounting and un-mounting.
- Vision, laser distance, and security systems should be intrinsic to the tool.

Objectives were achieved, allowing the tool to benefit from the advantages of repeatability, reliability, and precision provided by any robot. Conceptual design can be seen in Figure 3.

The final designed tool is composed by two complementary systems: a material application system (mechanical subsystems and actuators), and a vision and security system (camera, laser distance sensor, and security micro-switches).

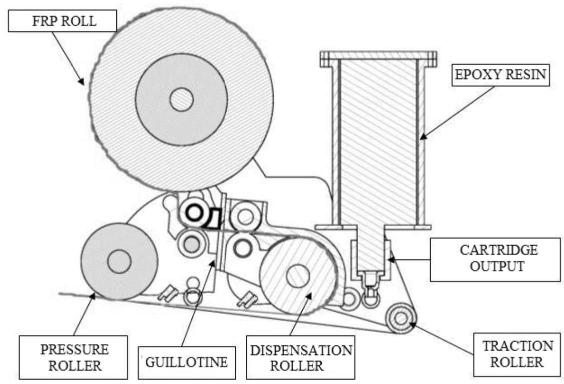


Figure 3. Tool's Basic Conceptual Design

Compressed air blowing was the option chosen for superficial preparation. Sand abrasion or hydro-demolition would have made sand or water available too, thus reducing system flexibility. The availability of compressed air is achieved with an on-board compressor mounted on the wheeled vehicle, though most tunnels also provide this service. Air flow is digitally controlled by system-mounted mini-electrovalves. Availability of compressed air also led us to one of the system's fundamental breakthroughs: Resin cartridges' push is triggered by canalization of compressed air towards its piston (Figure 4), therefore obviating the need of a motor, gear reductions, and additional mechanical components. This frees the tool from extra weight, size, and cost.

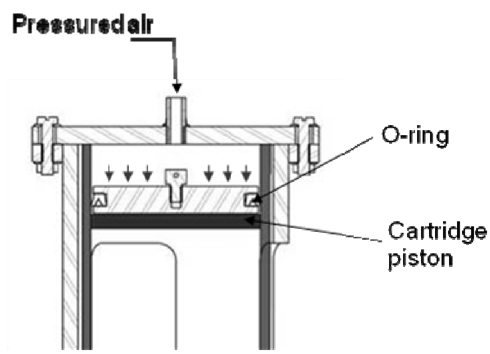


Figure 4. Resin Activation Mechanism

A sophisticated gear-and-roller-conjunction system was designed for the dispensation of FRP strips. Commercial FRP is distributed in the form of a long roll. In order to cut the roll into strips of the desired length, the gear-and-roller system is combined with a cutting mechanism that is activated by relative movements between the tool's internal mobile parts and the tunnel's surface. Linear actuators (pneumatic cylinders) were incorporated to automate the cutting process without having to depend on the robot or manipulator's strength against this surface. This is beneficial for the weaker range of manipulators. A linear dispenser, similar to the one used for air in the superficial preparation process, was designed to spread resin over the tunnel's weakened surface. Pre-impregnated FRP rolls were discarded due to possible undesired FRP-tool interaction due to adhesive properties. In the designed system, rollers, gears, and resin dispensation work simultaneously as the tool advances to obtain a compact resin-FRP-resin sandwich mounted on the tunnel's surface (Figure 5).

An Axis IP surveillance camera and an ultra-light precision Laser Telemeter sensor are the two main components of the vision system. Sensor ray and the camera' central pixel point are oriented parallel to the axis of the resin cartridge. Stereo vision was discarded due to its ineffectiveness on homogeneous patterns, such as tunnel concrete surfaces. The IP Surveillance Camera obtains images with a streaming resolution up to 1280x1024 pixels and 30 frames per second, sufficient for having real-time visual information in a control cabin. Micro-switches activated by contact with the tool's mobile parts assure mechanical safety.

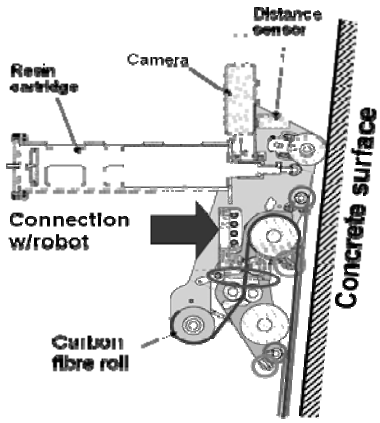


Figure 5. FRP Application Scheme

HMI and Control Software

Both HMI and control software were coded entirely in C++, a powerful multi-platform object-oriented programming language. The user-friendly HMI guides the system operator intuitively through the inspection and identification tasks (Figure 6), for superficial preparation and epoxy resin injection, or superficial preparation and FRP adhesion.

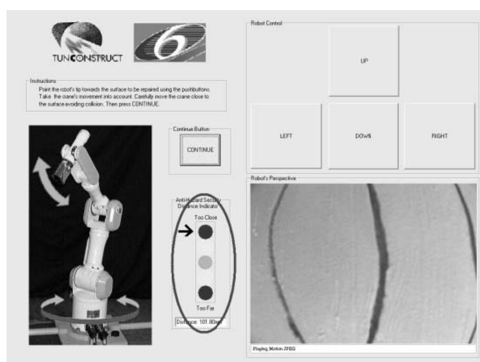


Figure 6. Intuitive HMI w/Distance Indicator

A virtual traffic light warns user if he or she is putting any part of the system's integrity in risk. (Figure 6, circled red). Both procedures converge into the final step: pressing a simple 'perform automatic operation' screen button. Previsualization of the desired process is also included in the software pack (Figure 7).

Figure 6. Intuitive Guided HMI Screenshots

Internally, software parameters are set in headers, and additional calibration information is calculated at run-time. Reusable spline and geometrical calculation libraries, developed specifically for this innovative robotic application, are integrated into the software. Barrel-type image distortion is corrected using the previously calculated intrinsic parameters of the camera. Communication schemes between HMI and control software are implemented over TCP/IP, the Internet Protocol. This means the operator could, technically, run the HMI software to identify fissures and weakened surfaces from any place in the world.

Laboratory and Field Tests

The RATIM was first tested indoors, using models manufactured by FRP and epoxy resin providers (Figure 8). Models were disposed in horizontal, vertical, and intermediate positions to assure material consistency (resin dripping would have been undesirable), robot configuration singularity avoidance [7], as well as correct robot-tool interaction (collision avoidance) [8].

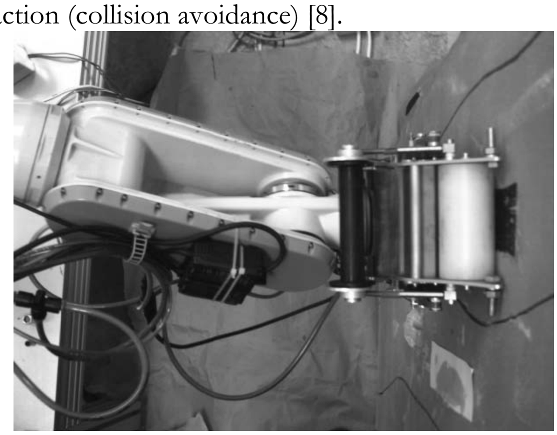


Figure 7. Laboratory Tests at UC3M

After having succeeded in proving effectiveness in laboratory tests, the complete integrated system was tested outdoors. Demonstrations were performed in real, non-controlled environments in tunnels in Bembibre, León (Figure 8).

Figure 8. RATIM System Tested Outdoors

There too, tests provided satisfactory results. Test samples were extracted from the tunnel's treated surfaces. Stress and strength parameters, contrasting equation (1) results, were analysed and validated at Sika S.A.U.'s laboratories.

Conclusions

The Robot-Aided Tunnel Inspection and Maintenance System is definitively the first system of its kind, a precedent for any robotic tunnel inspection and maintenance system of the future. It has been developed for inspection, maintenance, and reparation of small fissures and weakened surfaces. This fulfils the philosophy of true preventive maintenance. Reparation of small fissures beforehand will prevent the evolution of these small fissures into great, steep cracks. Just as well, reinforcement of weakened structures will prevent solicitations that can provoke the deterioration of the rest of the tunnel's infrastructure. Main characteristics of the RATIM system include reduction of operation times, a quick system and HMI learning curve, and a global increase of commodity and productivity for operator and society. Operation can be achieved, risk free, without need of stopping traffic flow or mounting colossal scaffolds.

References

[1] C. Balaguer (2000) Open issues and future possibilities in the EU construction automation IAARC ISARC '00, Taipei, Taiwan, K21-32.
[2] Nanni A.; Bradford N.M. (1995) FRP Jacketed Concrete under Uniaxial Compression, Construction and Building Materials, Vol. 9, No. 2, 115-124.
[3] H. Toutanji (1999) Stress-Strain Characteristics of Concrete Columns Externally Confined with Advanced Fibre Composite Sheets, ACI Materials Journal, Vol. 96, No. 3.
[4] M. R. Spoelstra; G. Monti (1999) FRP- Confined Concrete Model, Journal of Composites for Construction, Vol. 3, No. 3, 143-150.
[5] S. Lee; M. Gil; K. Lee; J. Lee; C. Han (2007) Design of a Ceiling Glass Installation Robot, IAARC ISARC '07, Vol 2.4, 247-252.
[6] A.Giménez; J.M.Pastor; V.M.Padrón; M.Abderrahim (2002) Unmanned Vehicles for Aerial, Ground and Naval Military Operations. Elsevier, NATO RTO series, No. 52, 161-166.
[7] C. Balaguer (2006) Robótica y Automatización, Fundación COTEC, ISBN: 84-95336-61-8.
[8] L. Moreno, S. Garrido, C. Balaguer (2003) Ingeniería de Control. Modelado y Control de Sistemas Dinámicos. Ariel, ISBN: 84-344-8055-7.

A Data Fusion Model for Location Estimation in Construction

S.N.Razavi[1] and C.T.Hass[2]

[1] PhD Candidate, Department of Civil and Environmental Engineering, University of Waterloo, 200 University Ave. W., Waterloo, ON, N2L 2V3; PH (519) 888-4567 Ext. 33929; FAX (519) 888-4300; Email: snavabza@engmail.uwaterloo.ca

[2] Professor, Director of the Centre for Pavement and Transportation Technology, Department of Civil and Environmental Engineering, University of Waterloo, 200 University Ave. W., Waterloo, ON, N2L 2V3; PH (519) 888-4567 Ext. 35492; FAX (519) 888-4300; Email: chaas@civmail.uwaterloo.ca

Abstract

Materials tracking is a key element in a construction materials management system. Deploying a cost-effective, scalable, and easy to implement materials location sensing system in real world construction sites has very recently become technically and economically feasible. The evident drawback of the current cost-effective and scalable systems is lack of accuracy and robustness.

In this research a data fusion model is used on an integrated solution for automated identification and location estimation of construction materials, equipment, and tools. Data fusion is intended to increase confidence, achieve better performance for location estimation, and add robustness to operational performance. The proposed model is a modified functional data fusion model for the application of construction resource location estimation and is based on the US Joint Directors of Laboratories (JDL) model. The paper presents some preliminary and promising results of applying the fusion model on construction field trial data.

Introduction

Material tracking is a key element in a construction materials management system. The unavailability of construction materials at the right place and at the right time has been recognized as having a major negative impact on productivity. Reducing unsuccessful searches for such materials would reduce wasted supervisory time, crew idle time, and disruptions to short interval planning. Conversely, understanding the materials flow over time helps to increase labor productivity, reduce materials stock piles, and reduce materials management manpower.

In an initial attempt to automate materials tracking Caldas et al. (2006) implemented a GPS and hand held GIS based mapping approach that demonstrated some promise for time savings and reduced materials losses under certain conditions. More sophisticated and automated, wireless sensor network based, data collection technologies, using GPS and RFID (Radio Frequency Identification), are being developed for a wide spectrum of applications. Specifically more recent research is demonstrating that coupled with mobile computers, data collection technologies and sensors can provide a cost-effective, scalable, and easy to implement materials location sensing system in real world construction sites (Akinci 2002; Song 2006a; Caldas 2006; Grau 2007, Teizer 2007). The evident drawback of the current cost-effective and scalable systems is lack of accuracy and robustness.

To address these problems, this study incorporates a framework for an integrated solution for automated identification and localization of construction materials, equipment, and tools for large industrial construction projects. A critical element of this framework is the location estimation problem in particular. Therefore, developing a data fusion method for location estimation that is robust to measurement noise while having a reasonable implementation cost would be advantageous. Fusing the different sources of location data is intended to increase confidence, achieve better performance for location estimation, and add robustness to operational performance.

In this framework, a range of simple to complex sensors can be utilized such as RFID transponders, GPS receivers, RFID readers, RFID with GPS chips, ultrasound, infrared and others. It is assumed that a small subset of sensors will have a priori information about their locations. This may happen because they have been coupled with GPS receivers or GPS chips or because they have been installed at some fixed points with known coordinates. This subset is small because no matter how a priori location information is achieved, it is on average one or two magnitudes more expensive per sensor node than estimated location information. For example, many geomatics solutions exist for tracking items accurately and in real time but at a cost that is prohibitive for the problem described here. In addition, even sophisticated and expensive solutions experience multipath, dead space and environmentally related interferences to some extent. Thus, developing a method for location estimation that is robust to measurement noise while having a reasonable implementation cost is a challenge.

This paper is organized in different sections as follows. Data fusion concepts and models are introduced briefly in the next section to provide some background information to the readers. It follows by presenting a data fusion model for location estimation in construction. The field experiments conducted to obtain the experimental data is presented next. The paper provides some preliminary and promising results of applying the fusion model on the construction field trial data.

Background

Data Fusion

Data fusion is a process of combining data or information to estimate the state of an entity. More often, the state of an entity is referred to as a physical state like identity, location, motion over a period of time and others. The human brain can be considered the best example of a data fusion machine.

Functional, process and formal models are three different categories of data fusion models (Steinberg 2001). A functional model can show the primary functions, relevant databases and the interconnectivity among the elements. A functional model does not show a process flow within a system. This means that levels in a functional fusion model should not necessarily perform sequentially. The US Joint Directors of Laboratories (JDL) model is an example of the functional model. Fusion researchers can develop their own models or adopt one of the existing models. Fusion of data results in many quantitative and qualitative benefits.

Building Information Modeling (BIM)

Building Information Modeling (BIM) is an approach to design, construction, and facility management in which a virtual model of a building is constructed digitally. The model contains precise geometry, spatial, and temporal relationships, 3D geographic information, and quantities and properties of building components to support construction, fabrication, and procurement activities and modeling of the building lifecycle (Eastman 2008). BIM can also be integrated with Cost and Schedule Control and Other Management Functions. It can be used to demonstrate the entire building lifecycle including all stages of building, and it is a method for sharing information. It may also ease communication between architects, engineers and construction professionals (Elvin 2007). Usually it is implemented in the form of a standard, and it is related to BrIM (Bridge Information Modeling) and other similar models.

Multi Level Data Fusion Model for Location Estimation in Construction

Model Architecture

Figure 4 describes a modified functional data fusion model for the application of construction resource location estimation. It is based on the JDL model because it is the most widely used system for classifying the data fusion based functions. The first two levels are called low level data fusion and the second two the high level fusion steps and the last level is called a meta-process. In the following figure, the architecture, the data flow and the interrelationships among the fusion levels are illustrated.

The data sources for this model include:

- Different physical sensors
- Different location estimation algorithms
- Context:
    - Received Signal Strength Indicator (RSSI)
    - Positional Dilution Of Precision (PDOP)

- Time
- BIM
    - Georeferenced site map/layout and Drawings
    - Georeferenced 3D models
    - Environmental conditions
    - Schedule (not in the scope of this study)
    - As-builts (not in the scope of this study)
    - Procurement details (not in the scope of this study)

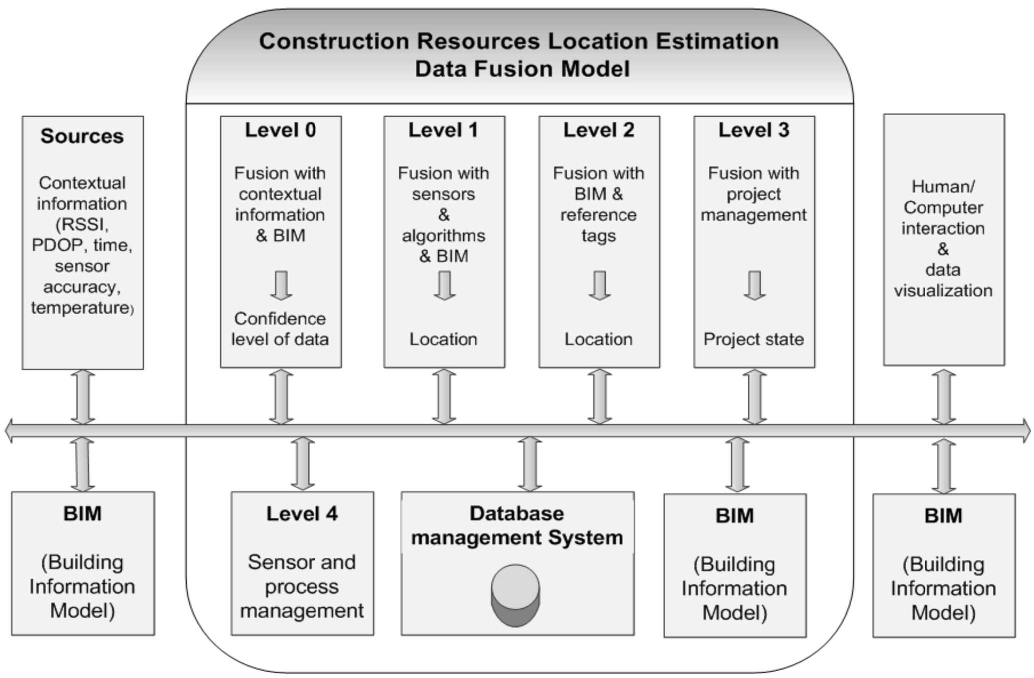


Figure 1: Data fusion model for construction resource location estimation

Data Fusion Level 0

Sensor data reliability assessment is the focus of interest in level 0. In other words, we want the sensor fusion system to utilize a combination mechanism so that the different sensors can "properly" contribute to the "results" in the information space. The "results" means the output of this fusion level and "properly" means with an established confidence, reliability or validity of the information.

Utilizing some location sensing technologies such as RFID, GPS, Ultrasound, infrared, and others gives us some rough location estimations of the materials on site. This very rough estimation is considered as a "read" event or a location observation. In this level, we focus on finding the confidence level of this location observation, based on the reliability and accuracy of the sensors at the time of observation and other layout contextual information that are available or can be adopted from BIM.

Different sensors have different accuracy and reliability factors that differ from each other and there is no simple solution for a proper combination of sensors. Combining the contextual information about the sensors and some other available context about the site layout is a reasonable means to obtain the confidence level of the observed location data. Because some of the context might not be available at all times or for all the sites, using this information is optional in the described solution.

A fuzzy inference system is used for this fusion level, with the ability of employing the contextual data according to their availability. This fuzzy system needs to be re-engineered for any new set of utilized sensors. A thorough description of this fuzzy level 0 fusion is presented in (Razavi 2008) for a scenario of RFID and GPS sensors.

Fuzzy representations and an inference system help to define the observation validity or "trustiness" more precisely. In this regard, observations are not valid or invalid anymore, but they have a degree of trust in the range of valid and invalid (Caron 2004). In other words, the confidence level of the observed location is the output of this fuzzy system that will be used to weight the fusion in the next level of the fusion architecture described here.

Data Fusion Level 1

Level 1 data fusion estimates the location of the construction resource using different reads, sensors, and location sensing algorithms. A fusion of different algorithms to get a more robust estimation can also fit into this level. Site layout and material membership to different site areas can be fused at this level. The Dempster-Shafer theory that is also known as the "theory of belief" or "theory of plausibility" or "evidential theory" is the primary method that has been used in this level.

In the current approach, when an RFID reader reads a tag, the combination of GPS/RFID data gives information about the location of the tag which is a hypothesis. This information can be modeled by a basic belief assignment because of the uncertainty in RFID read range due to the surrounding environment. To deal with this uncertainty, different beliefs are assigned to different subsets of cells centered on the GPS/RFID sensor set such that the sum of the all beliefs are equal to one.

In the simplest scenario, due to environmental and other factors, GPS and RFID are having different reliabilities for each event of "read". Therefore different read events can be considered as the independent observations that can be fused by the Dempster-Shafer theory.

Outputs of the RFID-GPS-based prototype were used as the inputs for the developed Dempster-Shafer-based algorithm in this fusion level. The prototype outputs were estimated locations of the tags, based on the observed read events for each tag. These estimated locations were calculated using centroid model. The following figure illustrates the hierarchical relationship among tag read events, estimated locations of the prototype, and the Dempster-Shafer-based fused estimation (Figure 2).

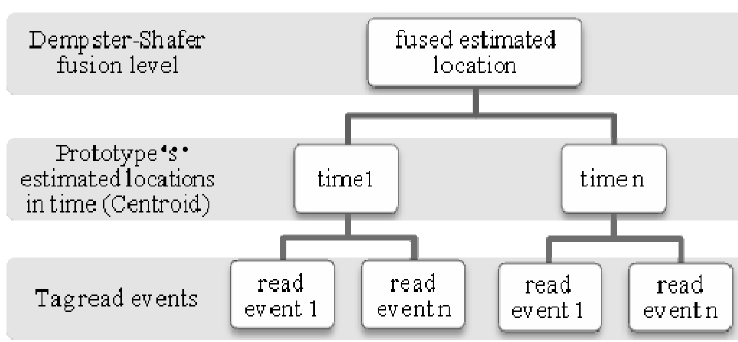


Figure 2: Hierarchical relationship representation among read events, estimated locations, and the estimations of the Dempster-Shafer fusion level

Data Fusion Level 2

It assesses the situation state by integrating the resource location information (Level 1 output) with contextual information, integrated BIM and/or other sensor data –LADAR, ultrasound or 3D Laser Scanner. The relationship between different construction resources and the site layout, as-builts and even schedule can be extracted based on the results of this level. This fusion level can result in a spatial/temporal relationship of elements and the building life cycle.

Fusion Level 2 is situation assessment on the basis of inferred relations among entities. Depending on the different physical and contextual information of the employed construction material locating approach, different solutions and techniques can function in this fusion level. For our approach, landmarks are used to assess the precision and correct the estimated locations of the target tags. The idea of using reference tags as some landmarks to adjust the estimated locations of the target tags is a feasible operation in this level of fusion.

In this framework, a cost-effective, arbitrary set of simple RFID transponders in some fixed and known positions in the construction site is utilized to possibly add accuracy to the estimated locations in the fusion level 1 of our target tags. As the reader agent is roving around and collecting the target tags data and estimating their locations through the level 0 and 1 fusion steps, reference tags data are also being captured at the same time and their locations would be calculated.

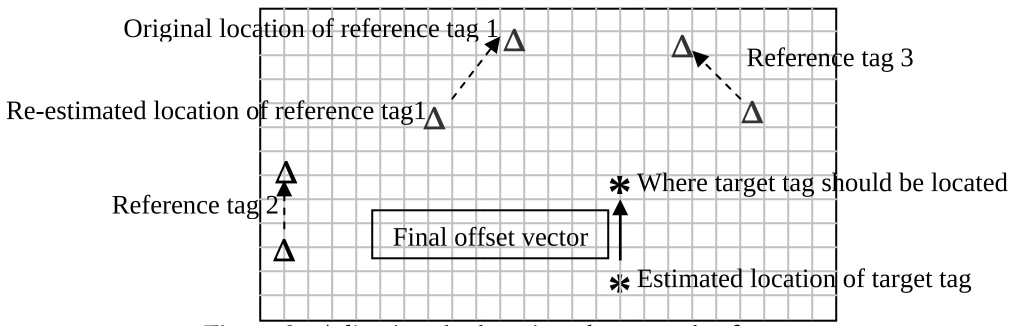


Figure 3: Adjusting the locations by several reference tags

The basic idea is using the vector of difference between pre-defined and re-estimated locations of the reference points and using this vector to offset the newly estimated target tag locations. The accuracy should improve if more than one reference tag can be employed in the framework. The composition of all the reference tags' offset vectors forms the final resultant offset vector (Figure 3).

Data Fusion Levels 3, 4 and Human/Computer Interaction

Level 3 is estimating the project state. This level involves integration with the project management system and is out of the scope of the current work. Level 4 improves the results of the fusion by continuously monitoring and assessing the sensors and the process itself. We may also evaluate the need for additional contextual information or sensors in this level. The need for calibrating the sensors or modifying the process may be assessed in this level. Human/Computer interaction can also be summarized in a data visualization and navigation module as well.

Conducted Field Trials

Field trials were conducted to obtain experimental data to validate the data fusion model and to demonstrate the feasibility of employing the components, methods and technologies developed. A large industrial construction project in Toronto hosted one field trial. An RFID-GPS-based location estimation prototype was used to conduct a comprehensive series of experiments with 375 tags to test the feasibility of tracking and locating some critical components on a construction site and its supply chain.

The data for testing the model are the coordinates of each tag ID on the lay down yards that have been logged on a daily basis for more than five months since the final RFID utilization started on the job site in August 2007. The estimated size of the data set is 100 days of data logging multiplied by average 100 tags on the site per day multiplied by a typical dozen reads per tag per day (Razavi 2008).

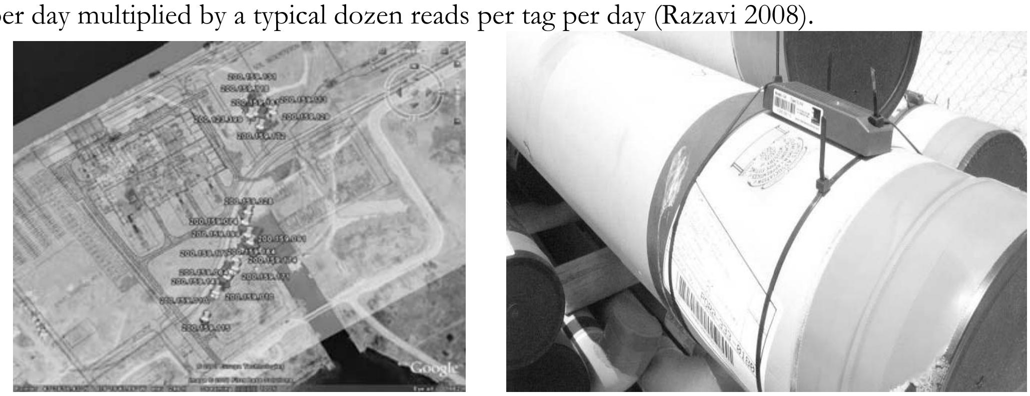


Figure 4: Sample map including some RFID tag location (left), and a sample tagged item (right)

The daily location data is saved in the format of .kml to be opened in the Google Earth map environment for visualizing the location information. An AutoCAD drawing of the site plan that was overlaid on the Google Earth aerial photo provided more landmark reference details for the locations on the site. Maps created in different granularity and various scales to allow proper visualization by field workers. Figure 4 presents a tagged item and a sample map.

Preliminary Results

Data on about 10,000 tag locations is available which represents an average of 100 location estimates for each tag in the field trial period. A case study of a sample data subset – belongs to the site warehouse – is presented in this section. For the subset of data used in this paper, the tags' location data were logged by GPS-enabled readers for 109 tags, three times per day, for four consequent days. The following figure represents the data distribution with respect to the distance between the estimated location – of the prototype based on the reads – and the real location of a tag.

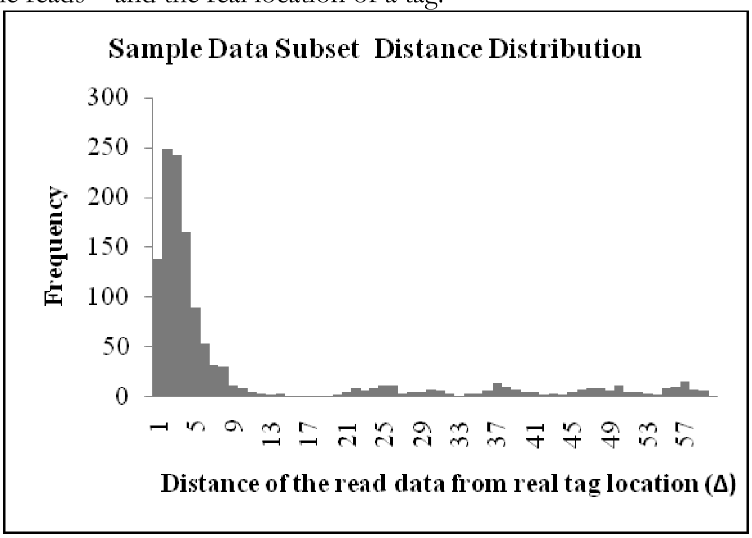


Figure 5: The distribution of the distance for the sample data subset

RFID read rates were sporadic, ranging from ten reads of a tag per minute to periods of hours without reads. Figure 6 presents a case study on how real-time fusion of two algorithms – Dempster Shafer and Centorid- can result in a more accurate estimation of location. In this case study, 8 read events of an RFID tag has been introduced sequentially to the fusion algorithm which represents the fusion level 1 in the implemented model. The final estimated location is equal to the center of the darkest blue area that corresponds to the highest pignistic probability.

Conclusions and Further Research

A functional model was presented for data fusion for location estimation of RFID tagged materials on a construction project. A fusion of two sensors, GPS and RFID, and two algorithms, Dempster-Shafer and Centroid, have been investigated to assess the location for the fusion level 1. Promising preliminary results are presented.

Further results will be reported in the near future. The challenge is to fuse data from simple to complex sensor sources, and contextual information, to estimate object location for tens of thousands of construction objects at an adequate frequency and in a scalable manner. It is expected that integration of fusion levels 0 and 1 will demonstrate significant performance enhancement with respect to measurement noise and will be robust to future advances in technology.

Acknowledgement

The authors thank Carlos Caldas, Paul Goodrum, David Grau, Hassan Nasir, Duncan Young, Paul Murray, and the undergraduate students involved in the field experiments and broader study.

References

[1] Akinci, B., Patton, M., and Ergen, E. (2002). "Utilizing Radio Frequency Identification on Precast Concrete Components - Supplier's Perspective." the Nineteenth International Symposium on Automation and Robotics in Construction (ISARC 2002) September 23-25, 2002, Washington, DC USA, 381-386

[2] Eastman, C; Teicholz, P.; Sacks, R; and Liston, K. (2008). BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors, John Wiley and Sons, NY.

[3] Elvin, G. (2007). Integrated practice in architecture: mastering design-build, fast-track, and building information modeling, J. Wiley & Sons press.

[4] Caldas, C., Grau, D., and Haas, C. (2006). "Using Global Positioning Systems to Improve Materials Locating Processes on Industrial Projects." Journal of Construction Engineering and Management, 132 (7), 741-749.

[5] Grau Torrent, D. and Caldas, C.H., (2007). "Field Experiments of an Automated Materials Identification and Localization Model" Proceedings of the 2007 ASCE International Workshop on Computing in Civil Engineering, Pittsburgh, PA, pp. 689-696.

[6] Razavi, S.N., Young, D., Nasir, H., Haas, C., Caldas, C., Goodrum, P., (2008). "Field Trial of Automated Material Tracking in Construction", CSCE 2008 conference, Quebec, Canada, June 2008.

[7] Razavi, S.N., (2008), "Data Fusion for Location Estimation in Construction," PhD proposal, University of Waterloo, Waterloo, ON.

[8] Song, J., Haas, C., Caldas, C., Ergen, E., and Akinci, B. (2006a). Automating the task of tracking the delivery and receipt of fabricated pipe spools in industrial projects. Automation in Construction, Elsevier, 15 (2) 166-177.

[9] Steinberg, A.N., Bowman, C.L., (2001). "Revision to the JDL data fusion model", pp.2-1--2-18. in Hall, D.L. and Llias, J., "Handbook of Multisensor Fata Fusion," CRC Press.

[10] Teizer, J., Lao, D., and Sofer, M. (2007). Rapid Automated Monitoring of construction Site Activities using Ultra-Wideband, Proceedings of the 24th International Symposium on Automation and Robotics in Construction, Kochi, Kerala, India, September 19-21, 2007, pp. 23-28.

[11] White Jr., F.E. (1987). Data Fusion Lexicon, Joint Directors of Laboratories, Technical Panel for C3, Data Fusion sub-Panel, Naval Ocean Systems Center, San Diego.

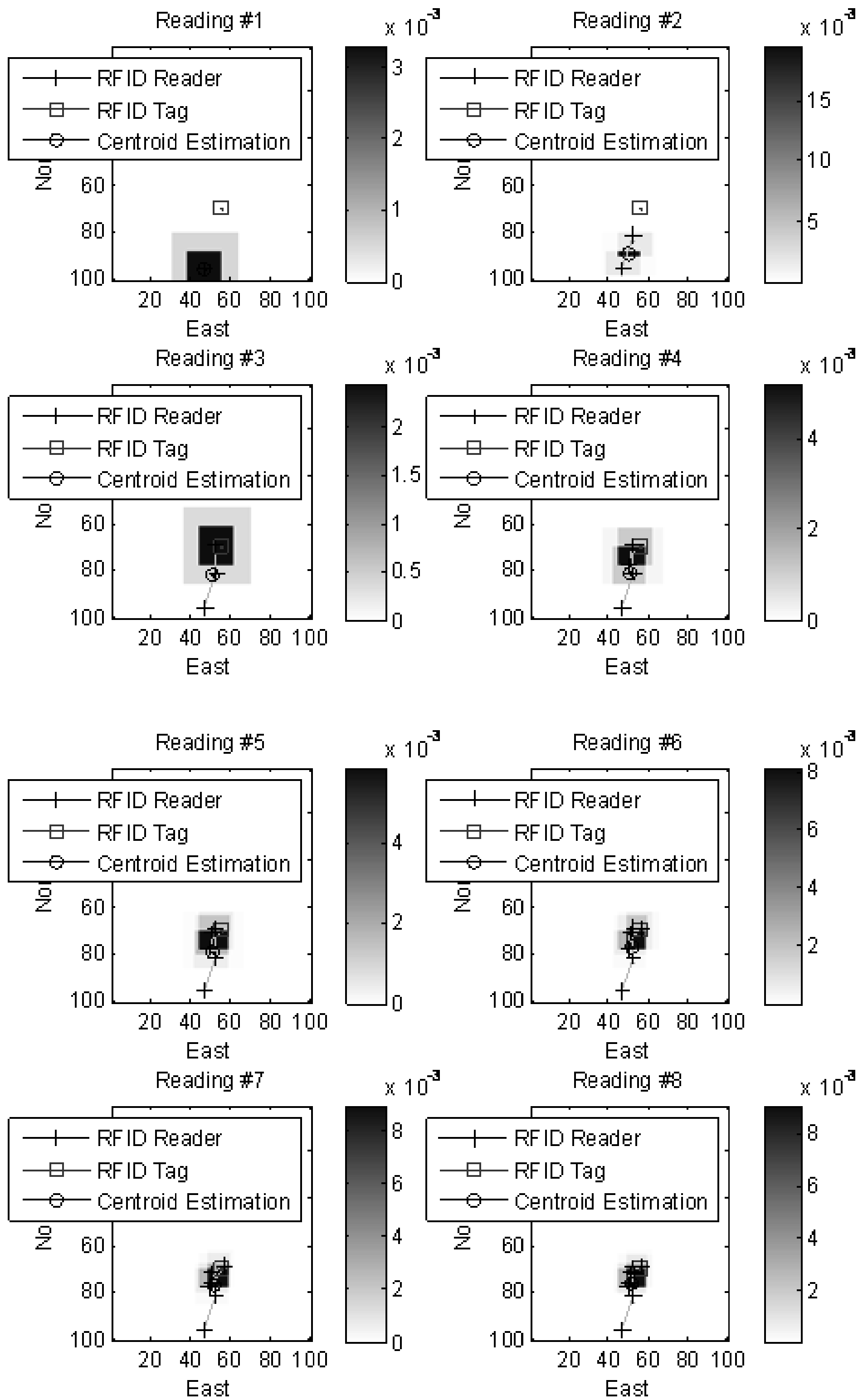


Figure 6: An Illustration on the fused Dempster-Shafer and Centroid methods for RFID tag ID of 200.159.095 after 8 Instances of Reading

Computational Models for Reduction of Quarry Waste for Breakwater Construction Projects

Venkateswara Rao Surati[1], Koshy Varghese[2] and K. Senthilnathan[3]

[1]M.Tech Student, Construction Technology & Management, Indian Institute of Technology Madras
[2]Professor, Department of Civil Engineering, Indian Institute of Technology Madras
[3]Head, Engineering Technology (Infrastructure), Larsen and Toubro Limited, ECC Division, Chennai

Abstract

Construction of Breakwater requires large volumes of appropriately graded rock. Geological characteristics and blasting practices at the quarry determine the gradation of the quarried rock. If the gradation of quarried rock does not meet design requirements, there is wastage of quarried rock or additional cost to reprocess it to meet requirements. This work utilizes mathematical models to forecast the gradation of rock from primary blasting. These models are programmed in the form of a spreadsheet to be used for decision support. By varying the blasting parameters the best fit between requirements and quarry yield can be found. In addition, a genetic algorithm based optimization model to determine the optimal values is also developed and illustrated with an example.

Keywords: Breakwater rock demand, Quarry yield, Blasting parameters, Spreadsheet, Genetic Algorithms.

Introduction

Breakwater construction involves quarrying, transportation and placing large volumes of appropriately graded rock. The large volumes of rock required and the gradation specified usually necessitates dedicated quarries for the project. Ideally, based on the geology of the quarry and blasting methods, the blasting pattern can be designed to ensure that the rock yield from primary blasting is close to the specified gradation. However, common practice is to produce large size rocks through primary blasting and then break these down to required specifications using secondary blasting. (Carlos, et. al. 1995) This process adds to the cost of the operation and results in considerable wastage of materials.

One of the key reasons for the current practice is that there are no standardized methods and decision support tools available to assist in forecasting the quarry yield for quarry characteristics and alternate blasting patterns (Clarke et. al. 2005). The objective of this work is to propose a decision support methodology based on available models and develop a tool to implement the methodology. The work utilizes the Rosin Rammler model (Vrijling et. al. 1990) to forecast quarry yield. A spreadsheet is used to encode the workflow of the methodology. In addition, the optimization features of the spreadsheet are used to automate the selection of the blasting parameters to minimize excess material. The utility of the tool is illustrated using an example.

Proposed Methodology

Figure 1 shows an overview of the proposed methodology. The design specifications for a particular section of the breakwater are considered as the initial input to the process. It is assumed that the design of the breakwater is frozen and the blasting parameters can be varied to ensure quarry yield obtained matches with the given design requirements.

The design requirement is based on the coastal parameters and properties of rock available in the quarry. The requirements will specify the various sections of the break water and the required gradation of rock for each section. Figure 2 shows typical sections of a breakwater and Figure 3 shows the rock demand requirements for the sections. The yield of the quarry must match with the demand to ensure adequate supply of materials with minimum wastage.

As shown in Figure 1, the yield of the quarry is based on geological characteristics (Rock Intact properties) of the quarry and blasting pattern utilized. Detailed mathematical models have been formulated to estimate the quarry yield.

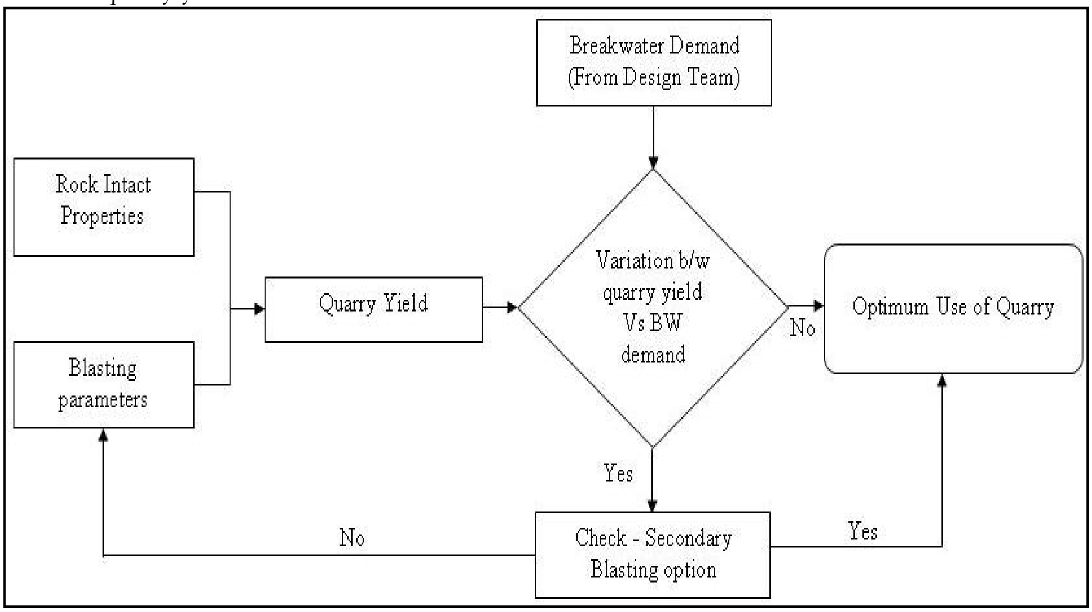


Figure 1 Proposed Methodology – Optimal use of Quarry

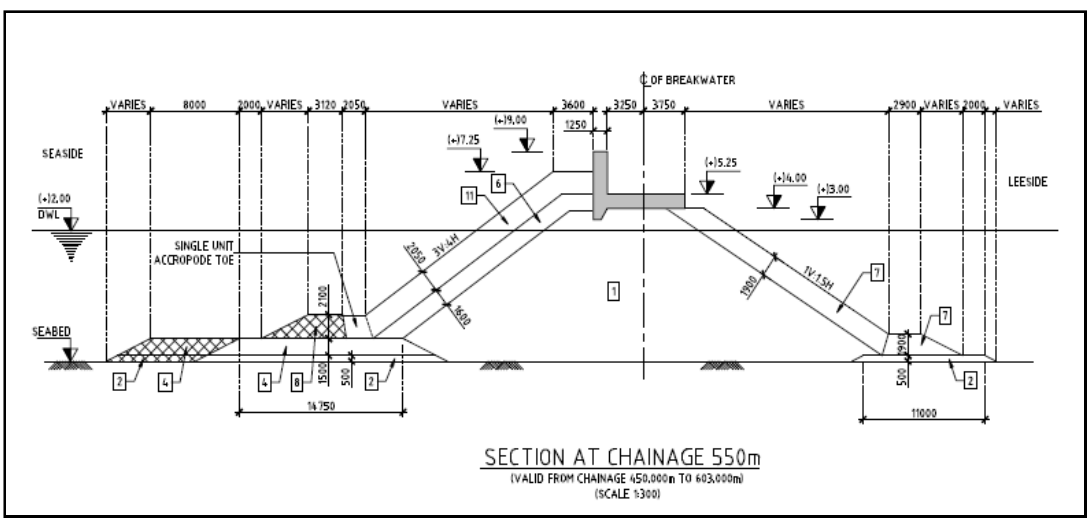


Figure 2 Typical Section of Rubble Mound Breakwater

Model for Quarry Yield:

The quarry yield for this work is calculated using the equation specified by equation (1): (Latham, et. al. 2006a) (Latham, et. al. 2006b)

$$\mathbf{Y = 1 - \exp\{-0.693(D_y/D_{b50})^{n_{RRD}}\}} \qquad \textbf{(1)}$$

Where:

D_y - *Specific particle size*

D_{b50} - *50% passing sieve size in the blast pile*

n_{RRD} - *Rosin–Rammler uniformity index for sizes*

D_{b50} is given by Kuznetsov equation; this suggests that average size is controlled by specific charge.

$$D_{b50} = 0.01. A . (V/Q)^{0.8} . Q^{0.167} . (E/115)^{0.633} \quad (2)$$

Where:

A = rock factor – calculated by equation (3)

Q = charge concentration per blast hole (kg)

V =volume of rock broken per blast hole (m^3)

E =relative weight strength of explosive (ANFO= 100, TNT=115);

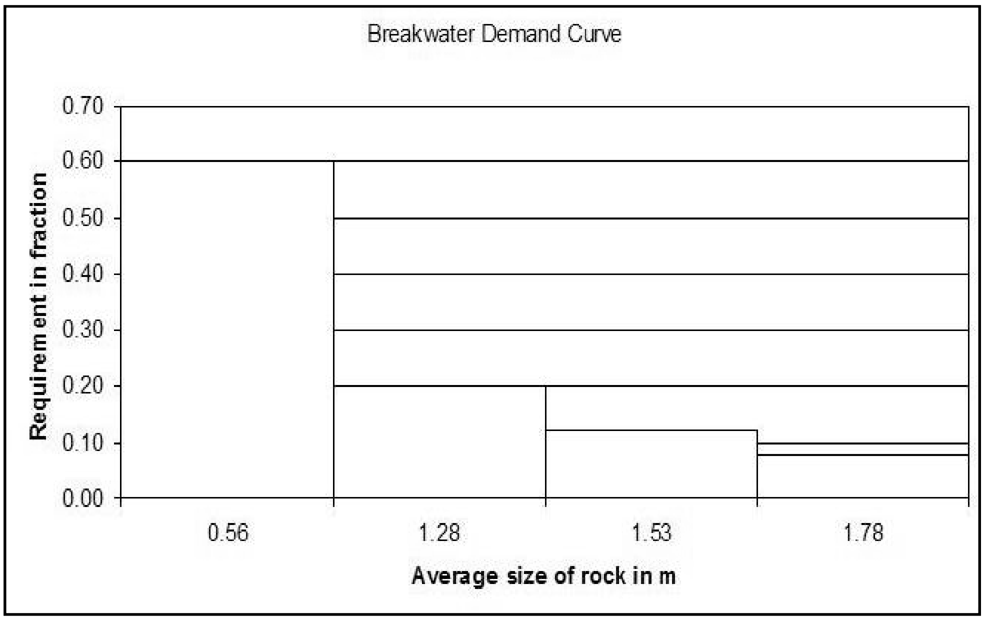


Figure 3 Demand Chart - Breakwater

Rock factor A:

$$A = 0.06(RMD + JF + RDI + HF) \quad (3)$$

Where

RMD (Rock mass description): 10 if powdery or friable, = JF if vertically jointed, 50 if massive rock

JF (Joint Factor): Joint Plane Spacing term (JPS) +Joint Plane Angle term (JPA)

JPS = 10 if average Principal Mean Spacing (PMS) <0.1 m; 20 if 0.1 < average PMS < to 1 m; 50 if average PMS>1 m.

JPA = 20 if dipping out of face, 30 if striking perpendicular to face, 40 if dipping into face

RDI = Rock Density Influence=0.025ρr (kg/m^3) −50

HF = Hardness factor=E/3 if E<50, or UCS/5 if>50, depending on uniaxial compressive strength UCS (MPa) or Young's Modulus E (GPa).

ρr = Rock Density

n_{RRD} in equation (1) is determined using Cunningham's uniformity index formula

$$n_{RRD} = X (2.2 - 14B/d)\{0.5(1+S/B)\}^{0.5}(1-W/B) \left(\left(|(BCL-CCL)|/L\right)+0.1\right)^{0.1} L/H \quad (4)$$

Where

d = borehole diameter (mm)

B = burden (m)

S = spacing (m)

BCL = bottom charge length (m)

CCL = column charge length (m)

L = total charge length above grade (m)

H = bench height or hole depth (m)

W = standard deviation of drilling error (m)

X = Design Pattern (Square pattern=1, Staggered pattern= 1.1)

Using the above models the yield of the quarry can be calculated. The yield of the quarry is expressed as the % of each rock grade range. The yield can be varied based on the following 4 key blasting parameters Charge Length, Burden, Spacing and Bench height. These are illustrated in Figure 4.

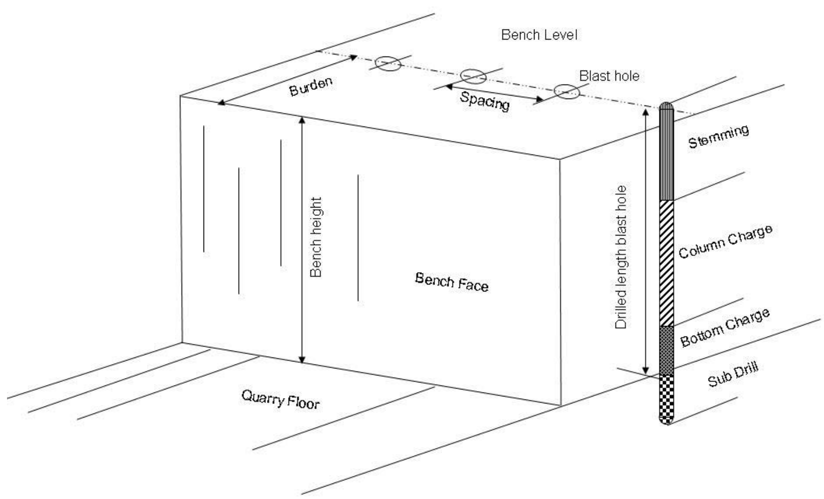


Figure 4 Quarry Blasting Parameters

The % of rock yield from the quarry in each fraction can be varied by altering the blasting patterns. Figure 5 shows three alternate yield curves and a comparison to the demand.

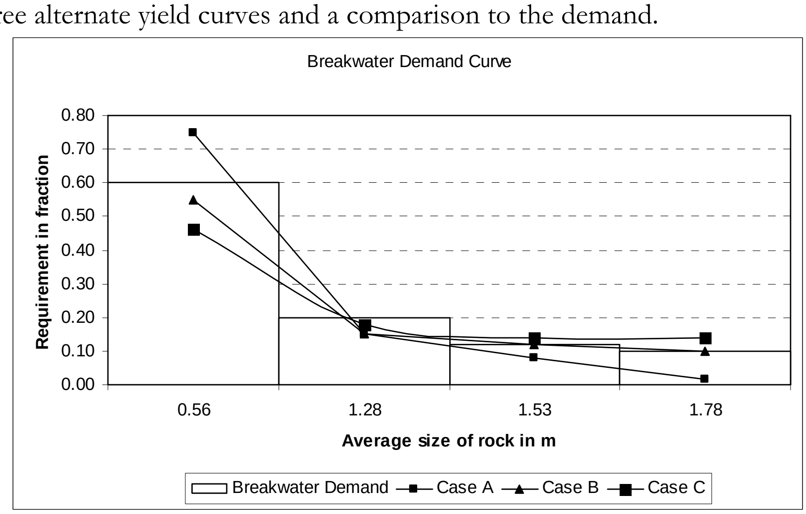


Figure 5 Alternate Quarry Yield vs Demand

In comparing the alternatives it can be seen that (i) For Case A, the yield of smaller size rock is more than the demand and larger rock is less than demand. This leads to wastage as the excess smaller rock cannot be used and additional quarrying has to be done to get adequate large size rock. (ii) For case B the yield of smaller size rock is less than demand and the yield of some fractions of larger rock are more than demand, in this case the larger rock can be broken up using secondary blasting to meet the requirements of the smaller rock. (iii) For case C, the yield of smaller rock, is much less than the demand and the yield of large rock is much higher, in such a case, extensive secondary blasting is required to break the excess large rock to meet the secondary rock requirements and this will be at a significant additional cost. Of the three cases presented, Case B is the preferred alternative as the primary yield is close to the demand and a minimal amount of secondary blasting is needed to meet the final demand requirements.

Using the models and methodology presented above a spreadsheet was developed to assist with the selection of the blasting parameters which can result in minimizing the wastage as well as cost of secondary blasting operation.

Matching quarry yield vs BW demand.xls

	B	C	D	E	F	G
24						
25	**Intact Rock Properties**		**Units**	**Pattern Design**		**Units**
26	Rock Type	Khodolite		Staggered or	1.10	
27	Rock Specific Gravity	2.65	SG	Hole Diameter	83.00	mm
28	Elastic Modulus	60	GPa	Charge Length	6.00	m
29	UCS	50	MPa	Burden	4.00	m
30	**Jointing**			Spacing	5.00	m
31	Spacing	0.2	m	Drill Accuracy SD	0.10	m
32	Dip	80	deg	Bench Height	10.00	
33	Dip Direction	1	deg	Face Dip Direction	0.00	deg
34	In-situ block	0.3	m	Powder Factor	0.06	kg/tonne
35	**Explosives**			Charge Density	0.15	kg/m^3
36	Density	0.9	SG	Charge Weight per hole	29.22	kg/hole
37	RWS	100%	(% ANFO)	**Fragmentation Target Parameters**		
38	Nominal VOD	4800	m/s	Oversize	1.78	m
39	Effective VOD	4800	m/s	Optimum	1.28	m
40	Explosive Strength	1		Undersize	0.56	m
41	**Predicted Fragmentation**			Blastability Index	7.16	
42	Percent Oversize	8.6%	m	Average Size of	54	cm
43	Percent In Range	39.4%	m	Uniformity	1.05	
44	Percent Undersize	52.0%	m	Characteristic Size	0.76	m

Blasting parameters / Design reqd_Second Blasting

Figure 6 Data Input Sheet

Spreadsheet Development and Usage

The spreadsheets developed to model the process are shown in Figure 6 and Figure 7. Figure 6 shows the sheet in which all the variables related to quarry properties and blasting are entered. The four decision variables are entered in this sheet (Burden, Charge Length, Spacing and Bench Height). Based on these variables the

% retention of each size is calculated and the estimates of quarry yield in % of each size are computed- this is displayed in Figure 7. The totally volume of rock has to be estimated based on the topography of the quarry and blasting pattern. The volume of rock yield for each size is calculated using the total volume and volume in each fraction. A comparison between the design rock requirement and the quarry yield gives the excess rock for that section.

The lower part of the second spreadsheet represents the secondary blasting in the form of a matrix. The rows & columns represent that size of rock. As smaller rock can be obtained from larger rock (but not vice versa) the cells above the diagonal represent the quantity of smaller rock which can be obtained from a particular row representing the larger size through secondary blasting.

To use the tool in a decision support mode the values of the decision parameters can be entered in the first sheet and based on the wastage obtained the decision maker can adjust the values until minimal wastage is obtained. An alternated usage is to be able to optimize the using an appropriate method.

Matching quarry yield vs BW demand.xls

	A	D	E	F	I	J	K	L
1	**Project:**	**Gangavaram Project (South Breakwater)**						
2	**Rock Specific Gravity**	**2.65**						
3	**Type of Rock**	**Khodolite**						
4	Class of Material	Average Size in m	Volume in Cum	Requirement in Fraction	Production in fraction	Production	Deficit	Over Production
5	Filter Layer/Core/Toe Armour (0 ~ 2 ton)	0.56	772,107	0.883	0.781	682,630	89,477	-
6	Secondary Armour (2 ~ 4 ton)	1.28	38,170	0.044	0.075	65,313	-	27,143
7	Primary Armour (4 ~ 6 ton)	1.53	44,798	0.051	0.037	32,255	12,543	-
8	Primary Armour (6 ~ 10 ton)	1.78	18,916	0.022	0.038	33,231	-	14,315
9	>10 ton	2.00			0.069	60,562	-	60,562
10		**Total =**	873,991	1.000	1.000	873,991		

12	Description of material	Filter Layer/Core/	Secondary Armour (2 ~ 4 ton)	Primary Armour (4 ~ 6 ton)	Primary Armour (4 ~ 6 ton)	Primary Armour (6 ~ 10 ton)	Waste
13	Filter Layer/Core/Toe Armour (0 ~ 2 ton)	-89477	27143	0	14315	48019	0
14	Secondary Armour (2 ~ 4 ton)		27143	0	0	0	0
15	Primary Armour (4 ~ 6 ton)			-12543	0	12543	0
16	Primary Armour (6 ~ 10 ton)				14315	0	0
17	>10 ton					60562	0

Blasting parameters / Design reqd_Second Blasting

Figure 7 Results Sheet

Optimization Model

In the manual mode a number of options will have to be explored to arrive at the optimal decision parameters. To automate this process the use of optimization was explored. The initial formulation objective was to minimize the variation between the requirement fraction and the production fraction. The constraints were to limit the fractions within a 5%-7 % range to ensure that primary blasting produced deficit of smaller rock and an excess of larger rock. The objective functions and constraints of the model are as follows:

Objective

$$\text{Minimize} \sum_{EachGrade} (\text{Produced Fraction - Required Fraction})$$

Constraints

% Deficit Production (0-2 Tons) <= 0.05

% Deficit Production (2-4 tons) <= 0.05

% Deficit Production (4-6 tons) <=0.05

% Excess Production (6-10 tons) <=0.07

% Excess Production (> 10 Tons) <=0.07

Initially the solution of the model was attempted with solver available in Excel. However, as there were logical statements in the spreadsheet model, the solver was not able to map the input parameters to outputs in continuous space. As a result the solutions obtained from these attempts were not valid.

A Genetic Algorithm (GA) can map input to output without the need for continuous space representation. The spreadsheet based GA - Evolver was used to represent the optimization model and solve it. Figure 8 shows the settings screen in Evolver.

The inputs were restricted to integers and the range of each input parameters was also specified. In addition to the objective of minimizing the excess rock, a model which required the deficit and excess rock to be balanced to archive a target value of 0 was also run. A population size of 50 was specified with the default crossover and mutation parameters. For the constraints specified the solution converged within 10 generations to the target value. Table 1 shows the values obtained for a case study of a recently completed breakwater project. The corresponding demand and yield graphs are shown in Figure 9.

Optimization was also attempted by narrowing the deficit and excess ranges in the constraints. This would provide a theoretically perfect answer; however, after numerous attempts no valid solutions were obtained.

The optimization model was primarily developed to investigate the applicability of modeling the demand yield problem and feasibility of generating results. Based on this preliminary study a more detailed optimization models can be developed. A natural extension is to consider the cost of secondary blasting also in the model and optimize both stages to ensure minimum wastage and cost of the operation.

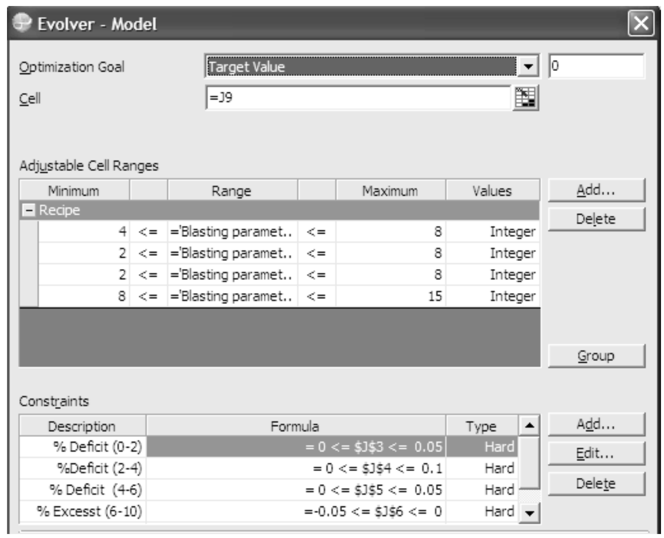


Figure 8 Genetic Algorithm Based Evolver Model

Summary

This study effectively illustrated how the efficiency of a complex task such as quarry production planning for breakwaters can be improved using appropriate quantitative models implemented in a spreadsheet. The quantitative models used have been validated on numerous quarries; the accuracy of the forecasts will depend on the applicability of these models to quarry under consideration. This field validation can be done during the trial blasting stage of quarry study. Further, the optimization offers potential to automate the decision process in selecting the appropriate blasting parameters. Continuing work in this area is focused on applying the models to live projects and further refining the optimization models.

Table 1. Optimization Results for a Breakwater Project

Class of Material	Required Fraction	Production Volume	Production	Deficit	Over Production
Filter Layer/Core/Toe Armor (0-2)	0.883	0.837	732,192	39,915	-
Secondary Armor (2-4)	0.044	0.042	36,916	1,254	-
Primary Armor (4-6)	0.051	0.021	18,933	25,865	-
Primary Armor (6-10)	0.022	0.024	20,837	-	1,921

>10 ton	0	0.070	65,113	-	65,113
TOTAL	1.000	1.000	873,991	**67,034**	**67,034**

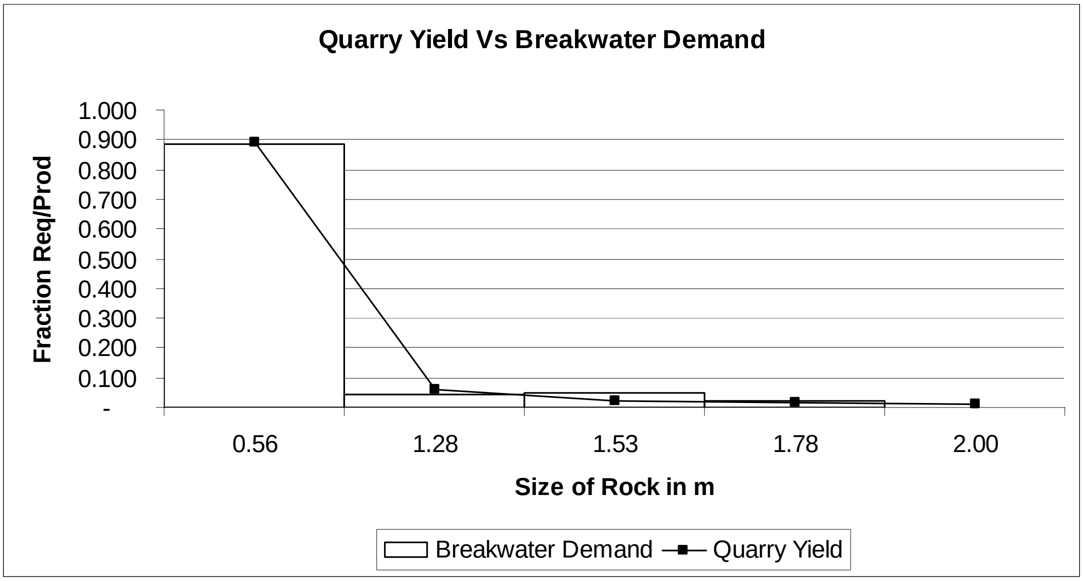


Figure 9 Yield vs. Demand Curve

References

[1] Carlos, L.J., Emilio, L.J., Francisco, J.A.C., and De Ramiro, Y.V (1995). "Controllable Parameters of Blasting", Drilling and Blasting of Rocks. P. 179-204.

[2] Clarke, N.J., Browning. N., Risbridger, P., and Drayton, N. (2005). "West Bay Coastal Defence and Harbour Improvements Scheme", Proceedings of the International Conference on Coastlines, Structures and Breakwaters, April 20-22. London, UK, p. 336-362.

[3] Latham, J.P., Meulen, J. V., and Dupray, S. (2006a). "Prediction of In Situ Block Size Distributions With Reference to Armourstone For Breakwaters", Engineering Geology, Volume 86, Issues 1-2, p.18-36.

[4] Latham, J.P., Meulen, J.V., and Dupray, S. (2006b). "Prediction of Fragmentation and Yield Curves With Reference to Armourstone Production", Engineering Geology, Volume 87, Issue 1, p.60-74.

[5] Vrijling, J.K., Nooy van der Kolff, A.H., (1990). "Quarry Yield and Breakwater Demand", Proceedings of the 6th Congress of the IAEG, Amsterdam. August 6-10.Balakema, Rotterdam, p. 2927–2933.

Integrating 5D Product Modelling to On-Site 3D Surveying of bridges

Teemu Kivimäki [1] , Rauno Heikkilä [2]

[1]Structural Design and Construction Technology Laboratory, Department of Mechanical Engineering, University of Oulu; Liikennetekniikka, Rakentajanaukio 4 A, 02150 TKK, Finland; PH +358 50 344 6465; e-mail: teemu.kivimaki@oulu.fi

[2]Structural Design and Construction Technology Laboratory, Department of Mechanical Engineering, University of Oulu; P.O.Box 4400 FIN-90014 Oulu, Finland; PH +1 530 219 7322; e-mail: rauno.heikkila@oulu.fi

Abstract

In this paper a new process integrating 5D product modelling and 3D on-site surveying for bridge construction is presented. This enables faster and less error prone surveying session preparations, fluent communication of design plans between survey teams, design office and other parties related to the construction project. A prototype system based on a total station and Tekla Structures CAD software is defined, and implemented by programming a .NET add-on to Tekla Structures and tested in field tests in an actually bridge construction project. Tests indicate prototype as a viable tool for surveying, but it still needs further development in usability and measuring features. Results are applicable also to building and road construction surveying.

Keywords: 5D, product modelling, surveying, on-site, real-time

1. Introduction

During a construction project, large amounts of information is generated, shared and exchanged. Studies have however shown that effective communication is hampered by lack of communication channels and processes.[1] One of these information-intense fields is surveying.

Today the state-of-the-art construction surveying tool is a total station. For surveying complex structures such as bridges, a total station is set up for a surveying session by uploading into its memory a specifically prepared simplified 3D surveying model. This model is built using coordinate information derived from printed 2D plans. An example of this kind of surveying model is shown in figure 1. This wireframe model is then used at the work site to mark the position of points to construction workers. The process of remodelling the structure for surveying every time there is a change in the plans is slow and error prone. Finally as-built models are rarely produced and information coming back to the designers from the construction site even during construction is usually only in form of phone conversations or excel sheets consisting of measured point coordinates.[2]

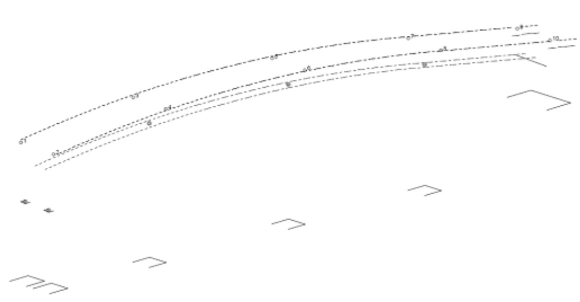


Figure 2. Current way of doing surveying planning: A 3D surveying model of the pilot pedestrian bridge modelled in 3D-Win CAD software, based on coordinates derived from printed 2D design plans.

Due to these factors, in the event of a required change in design, for example due to a prior error, information flows slowly and inconsistently between the different actors in a construction project, causing errors and delays in schedule.[3] To address these problems in the surveying process, research has been done on on-site surveying applications, mainly in the 1990's, and today there are even some commercial applications based on CAD software, namely Bentley MicroStation and Autodesk AutoCAD. None of these solutions have however become widespread, probably because in the past bridge design has been done in 2D, where the application of 3D on-site measuring is not possible. The emergence of Tekla Structures as a viable bridge product modeling software has enabled bridge designers to produce 5D-product models, which in turn enables 3D on-site measuring.[4]

The goal of this research was to define and test a new way of handling the surveying process by integrating 5D product models and on-site measurements. This would enable design plans and as-built information to be seamlessly communicated between designers, surveying teams and other construction site personnel.

2. Method

In this chapter, a new surveying process is defined, a prototype system implementing this process and field tests to test this prototype are described.

2.1 New surveying process

The designer models the bridge as a 5D product model. The model is then saved on a central server where all involved parties can access it via an internet connection. The surveying team can fetch this model directly from the server to the prototype surveying system described in chapter 2.2 and use it to do all surveying tasks required at the construction site. After surveying, the model is uploaded back to the central server where the designer and other involved parties, such as buyers, may immediately inspect surveying results (as-built model). Points that need to be surveyed can easily be marked in the model and change colour to reflect status. Deviations in as-built model relative to original plans are displayed as vector lines.

2.2 Prototype system

A Trimble 5605 DR total station and a Panasonic Toughbook CF-19 laptop computer suitable for field use were used as base components for the prototype system. A surveying software solution was needed to view and manipulate the 5D product model. Also a hardware solution for realizing data communication between the laptop and the total station was needed. Several software and hardware solutions were evaluated, and the following components were chosen to be realized in the prototype: Tekla Structures CAD as software solution and Satelline 3AS/d radio modems for realizing wireless serial port data communication. A surveying software was developed with Microsoft C#.NET language, that works as a user interface for the system and acts as a link between Tekla Structures and the total station's external command interface.

For the system to work, the bridge product model used must be modeled directly in the same coordinate system as is used at the construction site. The total station must be setup and oriented normally using the total stations basic setup process. When the total station and the bridge product model are in the same coordinate systems, measurements can be done.

2.2.1 Hardware

This section describes the prototype hardware. The main component was a Trimble 5605 DR total station, which was used to measure points selected from a bridge product model running on a laptop computer. Communication between the laptop computer and the total station was realized via serial port connection. The laptop could have been virtually any computer with a serial port and capable of running Tekla Structures CAD software, but for this prototype a Panasonic Toughbook CF-19 laptop was chosen because of its suitability for field use.

Since the surveying system was intended to be used by one person, and as the surveyor needs to hold the prism at the point being measured, the commands would have to be sent wirelessly to the total station from the position being measured. In the prototype system this was realized via Satelline 3AS/d radio modems, which are specifically designed to transmit serial data wirelessly. The system setup is visualized in Figure 2.

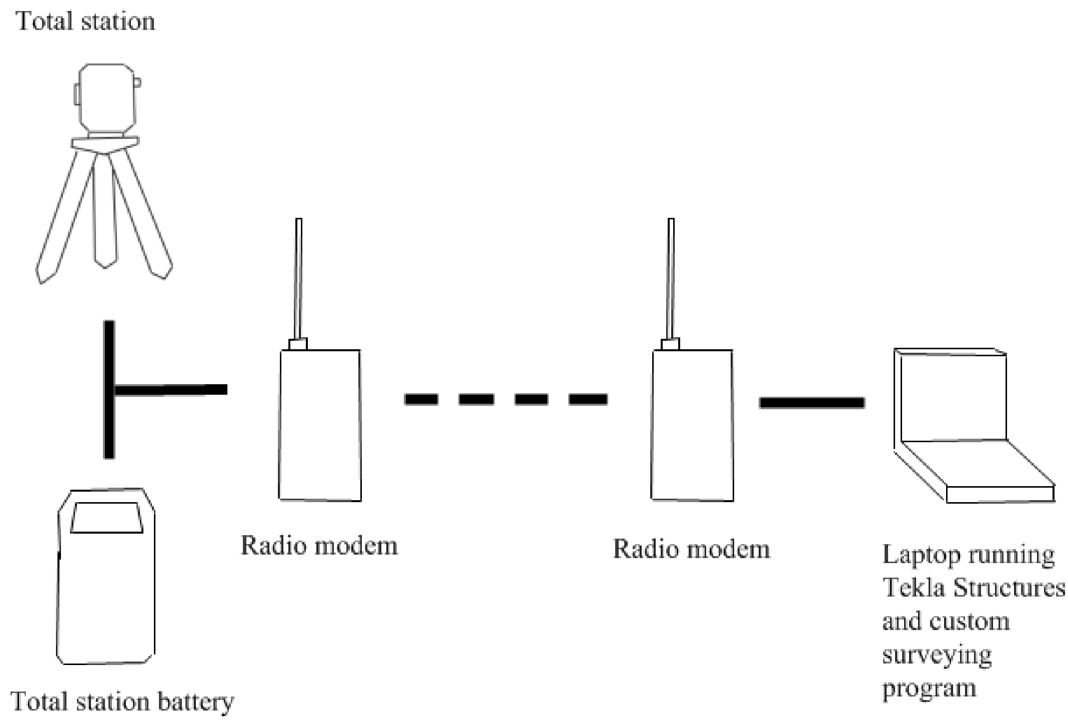


Figure 3. Prototype hardware setup

2.2.2 Software

The software solution was based on using Tekla Structures CAD and the programming interface it offers. A custom surveying software, a user interface, was needed to command the total station using only the laptop, basically acting as a link between Tekla Structures and the total station. Microsoft Visual Studio and C#.NET and Tekla Structures open programming API was used to realize a light software solution which communicated with Tekla Structures to allow the user to select points from the bridge product model using normal Tekla Structures user interface and snapping functionality. The commands were then executed on the total station by sending them via serial port connection, using the total stations serial port command interface.

The system can be used to guide the surveyor towards a selected point to mark it in the field for construction workers; the surveyor selects a point from the product model. The coordinates of this selected point are then displayed on the screen and the total station can then be commanded to measure and guide the surveyor to the selected coordinates. Alternatively the user may want to do an as-built measurement; select and point a feature, such as corner of a concrete slab, in the field. Then click the same feature in the product model, and the software adds a vector line displaying the difference in position, directly in the product model.

Figure 3 demonstrates a screenshot of the combined Tekla Structures and surveying software user interfaces.

2.3 Field tests

The prototype system was tested in a series of field tests at an actual bridge construction site. First test was conducted by measuring selected points from the 5D product model and comparing them with points marked in the field by traditional measuring methods. Second test was measuring points in a finished component and visualizing the data in the model (as-built). The product model was provided by the bridge designer in Tekla Structures format. The model is presented in figure 3. The prototype system and an operator during field tests are displayed in figure 4. Wireless transfer of product model via internet server between designer and survey team was not tested, but instead model was transferred by memory stick.

3. Result

Setting up and orienting the total station to the construction site coordinate system was straight forward. However since the bridge product model that was used had x- and y-coordinates reversed compared to the construction site coordinate system, problems occurred. After the problem was identified, it was solved finally by switching x- and y-coordinates manually in the surveying software user interface.

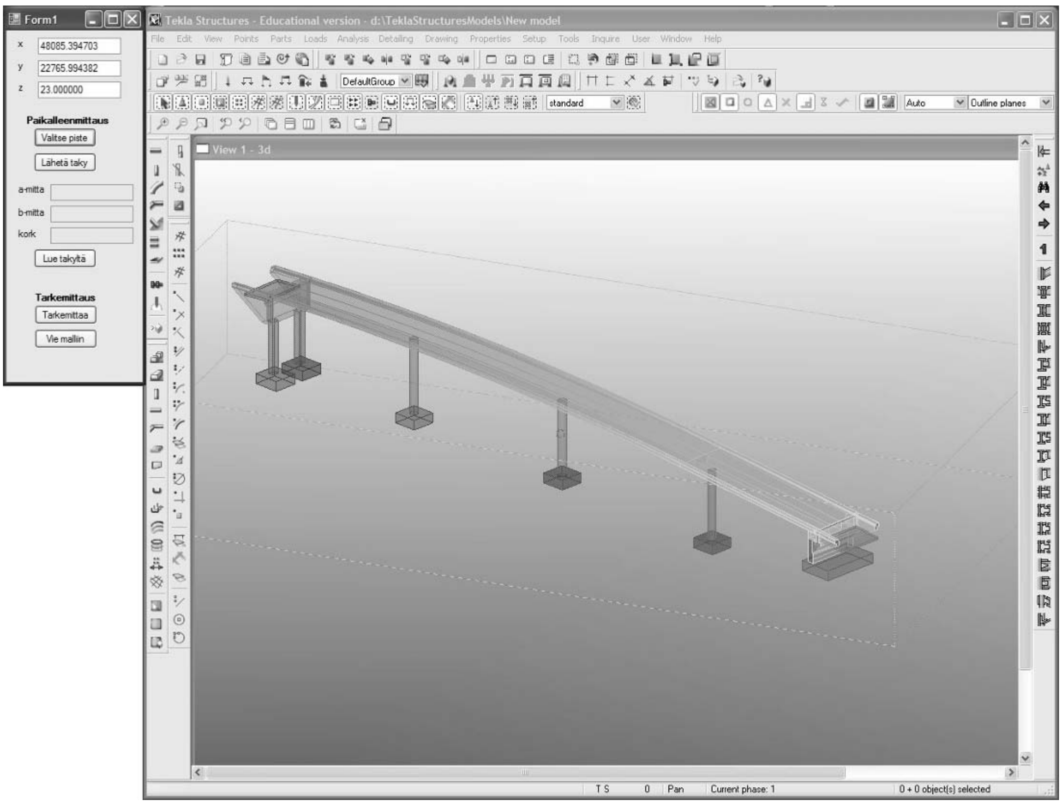


Figure 4. User interface of the surveying software (left side). Product model of the pilot pedestrian bridge is currently open in Tekla Structures (right side).

In the first field tests points previously measured by traditional methods were measured with the prototype system and yielded same results. This was of course the expected result, as only the method of writing and reading the data from the total station differs.

In the second test a finished component, bridge foundations, were measured. The actual foundations located at the construction site are shown in figure 5. The corners of the concrete cast were partially broken and posed a challenge on where to position the total station measuring prism. Finally all four corners were successfully measured and as-built vectors were immediately visible in the product model.

The as-built vectors generated in the product model are shown in figure 6.

4. Conclusion

The tests indicated that a 5D product model in Tekla Structures CAD software could be used directly to make on-site measurements.

As mentioned in chapter 1, several similar systems have been developed before, but using Tekla Structures and in-model vectors to indicate differences from planned positions is new. Tekla Structures also offers possibilities for sharing the different versions of the design plans with the different parties of the bridge construction project, over the internet. Optimal use of these plan sharing features would of course require wireless internet access at the construction site, which can be realized using different techniques, either by existing mobile phone GPRS or 3G network, or a WiFi network specifically setup for the construction work.

The cost effectiveness of a surveying process can significantly be improved by using position measuring directly from a 5D product model of the bridge, since survey preparation is faster, less labour intensive and error prone.

Figure 5. The prototype system at the field test site. Measured bridge foundations in the background (right).

Figure 6. Finished bridge foundations, partially covered in gravel.

The main challenge with the use of the system is that a surveyor often needs to move to difficult locations. A laptop is somewhat cumbersome to carry unless there is a specifically designed carrying harness. A visor type high quality display would solve this problem. The quality of current state-of-the-art visor displays is however not good enough.

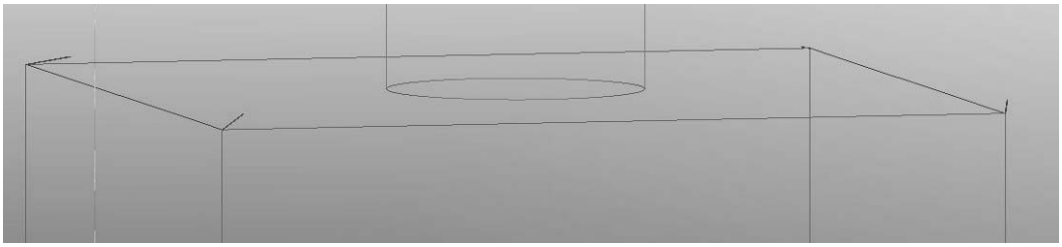


Figure 7. Vector lines in the product model displaying measured points relative to planned.

Further development is also needed in the surveying software user interface, to make it easier to use and also able to measure points relative to planes and lanes, which is a requirement in bridge surveying. Also road measurements are often done by the same survey teams as bridge measurements and they need to be done with the same surveying equipment. The results obtained are applicable also to building and road construction surveying.

Companies involved in funding this study have expressed interest in continued development of the prototype presented in this paper.

References

[1] Lee J, Bernold LE. Ubiquitous agent-based communication in construction. J Comp in Civ Engrg 2008 January/February 2008;22(1):31-9.

[2] Palkama M, Kivimäki T. Personal communication: Current surveying process in Destia and Finland in general. 2008.

[3] Beliveau Y. J. What can real-time positioning do for construction? Automation in Construction 1996;5(2):79.

[4] Heikkilä R, Karjalainen A, Pulkkinen P, Haapa-aho E, Jokinen M, Oinonen A, Jaakkola M. Siltojen 3D-suunnittelu- ja mittausprosessin kehittäminen ja käyttöönottaminen (Intelligent Bridge). Helsinki, Finland: Finnish Road Administration; 2005, ISBN 951-803-459-1.

Mining knowledge management strategies from the performance data of cop

Wen-der Yu[1], Shih-ting Lin[2], Shen-jung Liu[3] and Pei-lun Chang[4]

[1]Professor, Institute of Const. Mgmt., Chung Hua Univ., Taiwan, wenderyu@chu.edu.tw
[2]Master Student, Institute of Const. Mgmt., Chung Hua Univ., Taiwan, m09616015@cc.chu.edu.tw
[3]Assistant Vice President, CECI Engineering Consultants, Inc., Taiwan, sjliu@ceci.com.tw
[4]Engineer, Depart. of Business & Research, CECI Engineering Consultants, Inc., Taiwan, peilun@ceci.com.tw

Abstract

Knowledge community of practice (CoP) is a popular approach for knowledge management implementation in construction organizations including contractors and A/E firms. In order to evaluate and improve the performance of the knowledge CoPs, quantification methods for performance measurement were proposed in previous researches. Profound implications may be inferred from the performance data recorded from daily knowledge management activities. Such implications provide directions of valuable strategies for administration schemes and system modifications. To achieve such goals, the performance improvement patterns and rules should be identified. In this paper MS SQL Server® was adopted to performed Data Ming (DM) tasks that dig out the abovementioned patterns and rules from the CoP performance data. Three DM techniques (Decision trees, Clusters, and Association Rules) were employed to mine the rules and patterns existing in the 4,892 historic performance data recorded from the CoPs of a leading A/E consulting firm in Taiwan. Performance improvement strategies are then inferred and planned based on the rules and patterns discovered.

Keywords: **Knowledge Management; Data Mining; Consulting firm; Performance Measurement, Strategy planning.**

1. Introduction

Knowledge Management System (KMS) is a popular approach for knowledge management implementation in construction organizations including contraction and A/E firms. A KMS does not only provide a platform for knowledge generation, storing, retrieval, and sharing, but also enable an organization a tool to measure and monitor its intellectual property. In order to evaluate and improve the performance of the KMS, quantification methods for KMS performance were proposed in several previous works [1][2][3]. From those works, it was found that profound implications may be inferred from the performance data recorded in daily knowledge management activities. Such implications may indicate valuable strategies for increasing benefits resulted from the KMS both in terms of administration and system modification schemes. The key to achieve such objectives is finding out the performance improvement knowledge. The Data Mining (DM) and Knowledge Discovery in Databases (KDD) are proven to be very effective in mining patterns and rules residing in large databases [4][5][6][7].

In this paper, a case study is conducted on mining knowledge of improvement strategy from the performance data of a generic CoP in a leading A/E consulting firm in Taiwan, the CECI Engineering Consultants, Inc. (CECI). The proposed methodology combines two major elements: (1) a quantitative model for measurement of the performance of CoP; and (2) commercial DM software—Microsoft SQL Server®— for performing DM tasks. Totally 4,892 historic performance data were collected from nine selected CoPs of the case A/E firm for case study. Questionnaire surveys were conducted with the participants of CoP knowledge management (KM) activities via a web-based internet questionnaire surveying system. The survey results are then converted into data in the form acceptable for DM by the Microsoft SQL Server®. Three DM techniques (Decision trees, Clusters, and Association Rules) are employed to mine the rules and patterns existing in the performance data. Meaningful rules, useful patterns, and important association rules are found with DM. Performance improvement strategies are then inferred and planned.

The rest of this paper will be presented in the following manner: previous related works are reviewed in Section 2 to provide background of this paper; the methodology of knowledge discovery in CoP performance data is described in details in Section 3; then, a case study is conducted for mining of knowledge from CoP performance data of the case A/E firm; finally, findings from case are discussed and the concluded.

2. Review of related Previous Works

2.1 KMS in A/E Consulting Firms

Mezher et al. [8] reported a work on a KMS in a mechanical and industrial engineering consulting firm in middle-east. Their paper concluded the process of building a knowledge management system in the Mechanical and Industrial Department at DAR AL HANDASAH, which is a leading consulting firm in the Middle East and the world. Finally, the paper concluded the lessons learned from the experience of building the knowledge management system and the steps needed to improve it. Other works related to KMS in A/E firms were reported by the authors of this paper [1][2][3], which focus on a specialized KMS for emergent problem-solving, namely SOS, of a A/E consulting firm in Taiwan. Those works analyzed the characteristics of knowledge management (KM) activities in the A/E consulting industry and how KMS can improve the competitiveness of the firm.

2.2 Performance Measurement of KMS

The most related work in literature on performance evaluation of a KMS was a work done by del-Rey-Chamorro et al. [9] in Cambridge University. They developed an eight-step framework to create performance indicators for knowledge management solutions. del-Rey-Chamorro et al.'s work can be very useful for creating performance indicators of a KMS, however, their work was primarily developed based on the observations of KMS in manufacturing industry. Bassion et al. [10] addressed that in developing a conceptual framework for measuring business performance in construction should take into account the organization's business objectives. Yu et al. [3] proposed a quantitative model for measuring time, man-hour, and cost benefits resulted from a KMS of an engineering consulting firm. In the paper, details of the proposed quantitative KMS benefit models are presented with a case study application to an A/E consulting firm. It was reported from their study that the average time benefit (TB) is 63%; the average man-hour benefit (MHB) is 73.8%; and the average cost benefit is 86.6%.

A similar research was reported by Yu et al. [11] in quantifying the performance of KM activities of a generic CoP of an A/E firm. In their research, the KM activities of a generic CoP can be classified into two categories: (1) knowledge sharing activities; (2) problem-solving activities. Quantitative models were developed for both of the two types of KM activities.

3. Methodology of Knowledge Discovery in CoP Performance Data

3.1 General KDD Procedure

In this paper, the methodology of knowledge discovery in CoP performance data is based on the KDD procedure proposed by Han and Kamber [6] as depicted in Figure 1. A general process of KDD depicted in Figure 1 consists of the following detail steps: (1) Understanding the domain problem; (2) Extracting the target data set; (3) Data cleaning and pre-processing; (4) Data integration; (5) Data reduction and projection; (6) Choosing the function of data mining; (7) Choosing the data mining algorithm(s); (8) Data mining; (9) Interpretation; and (10) Using discovered knowledge—incorporating the discovered knowledge into the performance system, taking actions based on knowledge.

3.2 Data Mining (DM)

Data mining is an interdisciplinary field with a general goal of predicting outcomes and uncovering relationships in data [5]. DM is also the most critical step in the KDD process. While KDD refers to the overall process of turning low-level data into high-level knowledge, DM is the core mechanism that extracts useful knowledge from historical databases. It uses automated tools employing sophisticated algorithms to discover hidden patterns, associations, anomalies and/or structure from large amounts of data stored in data warehouses or other information repositories [12]. Data mining tasks can be descriptive, i.e., discovering interesting patterns describing the data, and predictive, i.e., predicting the behaviour of the model based on available data [13]. Data mining involves fitting models to or determining patterns from observed data. The fitted models can be viewed as inferred knowledge. In this paper, the Microsoft SQL Server® is adopted for DM tasks on the KMS performance data. Even though the Microsoft SQL Server® provides nine different DM algorithms (Classification, Estimation, Prediction, Association rule, Clustering, Sequential Pattern, Decision tree, Neural Network, Time series), three of them (Decision Tree, Cluster, and Association Rule) are selected for this research after testing with the performance data.

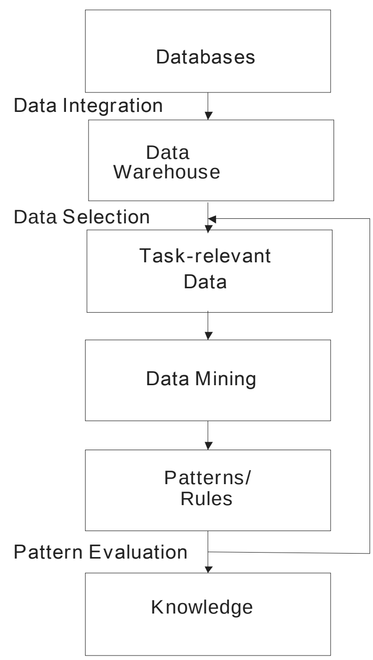


Figure 1 General process of KDD

3.3 Collection and Quantification of CoP Performance Data

A Knowledge Value Adding Model (KVAM) for quantitative performance measurement of KM activities in a CoP was proposed by Yu et al. [11]. The basic model of KVAM is depicted in Figure 2. There are two stages in a KM activity: (1) Raw knowledge creating process (RKCP) performed by the initiator of a KM activity; (2) Knowledge value adding process (KVAP)—performed by the participants/respondents of a topic in the CoP.

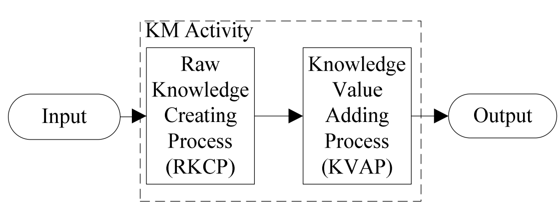


Figure 2 Basic model of KVAM [11]

It was propositioned by Yu et al. that "the amount of value of a KM activity is proportionate to the change of knowledge amount between Input and Output". Thus, the amount of Knowledge Value Added (KVA) of a KM activity can be measured by the difference between Input and Output of the process. This is calculated by multiplying the subjectively determined fuzzy terms of the two separate stages.

Two types of KM activities are identified: knowledge sharing and problem-solving. Fuzzy linguistic terms are defined to describe the RKCP and KVAP of the two types of KM activities. For knowledge sharing activities, five fuzzy terms are defined for RKCP. These fuzzy terms are called Raw Knowledge Creating Terms (RKCTs), including: (1) non-relevant; (2) data; (3) information; (4) knowledge; and (5) wisdom. Similarly, five fuzzy terms are defined for KVAP, Knowledge Value Adding Terms (KVATs), including: (1)

no-value-added; (2) get; (3) use; (4) learn; and (5) contribute. The similar definitions are given to fuzzy terms of the RKCP (RKCTs) and KVAP (KVATs), respectively, of the problem-solving activities. Details of the definitions for RKCTs and KVATs are referred to Yu et al. [11].The membership functions of the RKCTs and KVATs for both types of KM activities are assessed by questionnaires with the managers of the CoP. Techniques for determining the fuzzy means and sprays are used to construct the fuzzy membership functions [11].

4. Case Study

4.1 Background of Case A/E Firm and KMS

The CECI is one of leading A/E firms in Taiwan. It was established in 1969 primarily for the purpose of promoting Taiwan's technology and assisting in the economic development of Taiwan and other developing countries. The number of full-time staffs of the firm is about 1,700. Among those around 800 are in-house staffs in headquarters located in Taipei, the other 900 are allocated in branches and site offices around the island. Headquarters, braches, and site offices are connected by Intranet.

The structure of the case A/E firm consists of five business groups: (1) Civil Engineering Group; (2) Railway Engineering Group; (3) Electrical and Mechanical Engineering Group; (4) Construction Management Group; and (5) Business and Administration Group. Each business group includes several functional departments. The annual revenue of case A/E firm is around 4 billion TWD (128 million USD). According to the information disclosed by the firm, more than 1,700 A/E projects were finished in the past thirty years. Totally volume (construction budget) of the finished projects exceeds 300 billion USD.

4.2 Collection of CoP Performance Data

In order select appropriate CoPs with right cultural and enthusiasm in knowledge sharing, interviews were conducted by the research team during March 2007~May 2008 to meet with the managers of the CoPs. Finally, nine CoPs were selected: (1) Steel Community (associated with Structural Design Department); (2) Rail-Highway-Airport (associated with Transportation and Civil Department); (3) Supervision Art (associated with Construction Management Department); (4) Geotech (associated with Geotechnic Engineering Department); (5) Actuary (associated with Accounting Department); (6) Column-beam (associated with Structural Design Department); (7) Bridge (associated with Transportation Department); (8) Hydro-environ (associated with Hydraulic and Environmental Resources Department); and (9) Subway (associated with Railway and Mass Transportation Engineering Department).

Table 1 No. of surveyed KM cases

Sample statistics / CoP name	Total	
	Sampled	Valid
Steel Community	348	348
Rail-Highway-Airport	755	740
Supervision Art	674	663
Geotech	367	346
Actuary	182	177
Column-beam	111	107
Bridge	335	333
Hydro-environ	1876	1871
Subway	244	240
Total	4892	4825

The records of KM activities (including knowledge sharing and problem-solving activities) were collected from the nine CoPs from 2005/1 to 2007/12. Totally 4,892 KM cases were selected. Questionnaires were provided to the managers of the nine selected CoPs. See Table 1, among the 4,892 surveyed cases, 4,825 responses were valid with almost 99.9% of valid samples. Such high percentage of valid response was due to the effort of the research partner, CECI. An administrative mandate was ordered by the Business Research and Development to require all surveyed managers to participate in the questionnaire survey and respond timely. The data sets collected from questionnaire are then used for Kohonen learning to obtain the means of fuzzy membership functions associated with the associated RKCTs and KVATs.

4.3 Preparation of Data

The 4,892 historic CoP performance data were transformed into quantitative performance records by questionnaire surveys. Questionnaire surveys were conducted to the managers of the nine CoPs and all participants (initiators and responders) of the 4,825 KM records. After transformation, each KM activity performance record contains values of the seven attributes listed in Table 2.

Table 2 List of attributes in performance data

Item	Attribute name	Code	Description
1	CoP	ForumID	The name of CoP.
2	Article No.	TopicNo	Sequential No. of the article posted in a CoP. Totally, 4,892 articles are collected.
3	Technical code	TechCode	Classification code for the specialty of the articles.
4	Department	DeptCode	Department of the initiator. Totally, 41 departments.
5	RKCT	Kshare	Raw knowledge creating terms.
6	KVAT	Kapply	Knowledge value adding terms.
7	KVA	Kva	Knowledge-value-added calculated.

The Technical Code of Table 2 is a special classification system for all works and documents of the case firm. The Technical Code consists of three level: (1) Lifecycle code—first digit, describing the stage of the work in a project lifecycle; (2) Product & service code—digit 2 & 3, describing the product or service associated with the work; (3) function code—digit 4~6, describing the special function or technical domain of the work.

4.4 Performing Data Mining

The DM was performed with the Microsoft SQL Server® 2005. The general data DM procedure consists of the following 7 steps: (1) Adding a new analysis service project; (2) Adding a new data source; (3) Linking the new data source; (4) Setting up data view of the new data source; (5) Selecting data table and data view; (6) Completing adding new data view; (7) Selecting a DM algorithm for data mining. In this paper, all nine DM algorithms provided by the Microsoft SQL Server® 2005were tested with KMS performance data. Finally, three DM algorithms were adopted including Decision Tree, Clustering, and Association Rule.

4.5 Results of Data Mining

4.5.1 Decision trees

Procedure for mining of Decision Trees includes the following four steps: (1) Assigning dataset for DM to determine the variables for prediction; (2) Assessing and recognizing data type; (3) Entering DM structure name; (4) Decision Tree construction. After DM, six decision trees were constructed. Totally 13 meaningful rules were resulted from the Decision tree DM. The rules are named D1~D13. An example of the mined decision trees is shown in Figure 3.

Two examples of rules obtained from Decision Trees are: (1) D1, D2—If TechCode = "060N25" ("water pollution" technical domain) or "060N00" ("environmental" technical domain), then the Kshare is "non -relevant"; (2) D9—If DeptCode is "Architecture" and TechCode = "571C35" ("RC bridge"), then Kva is "extremely high" (>=93).

4.5.2 Clusters

Procedure for mining of Clusters includes the following seven steps: (1) Assigning dataset for DM to determine the variables for prediction; (2) Assessing and recognizing data type; (3) Setting up variables; (4) Setting up parameters; (5) Viewing relationships between variables and parameters; (6) Viewing Clustering models; (7) Visualizing final Clustering model. Totally 14 meaningful patterns were identified from the Clustering DM. The patterns are named C1~C14. An example of the mined clusters is shown in Figure 4.

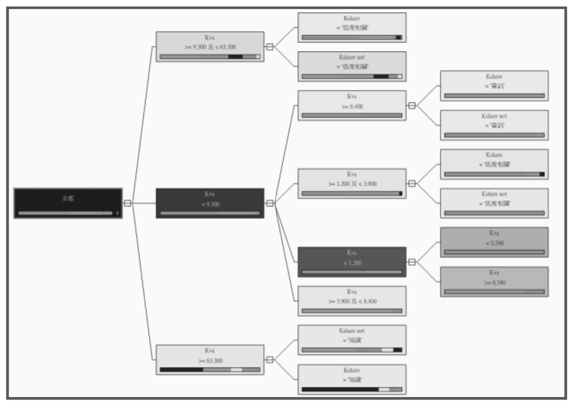


Figure 3 Example of mined decision tree

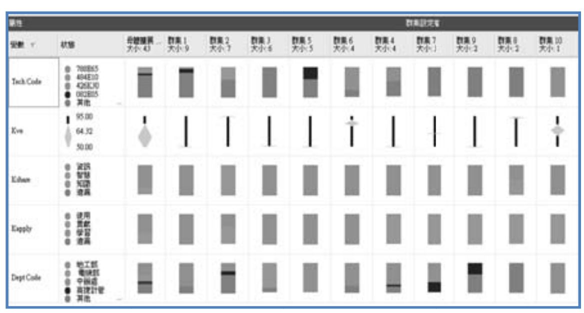


Figure 4 Viewer of mined clusters

Examples of rules obtained from Clusters are: (1) C1, 2, 4—low Kva clusters (Average Kva < 30) are identified as ForumID is "Hydro-environ" and DeptCode is "Hydraulic and Environmental Resources"; (2) C3—a high Kva cluster (Average Kva =83.26) is identified as ForumID is "Hydro-environ" and DeptCode is "Lioduey Supervision Office (a branch office of CECI in south Taiwan)"; (3) C8—a extremely high Kva cluster (Average Kva > 90) is identified as TechCode is "Foundation Engineering" and DeptCode is "Kaohsiung MRT Project Management"; (4) C10—a extremely high Kva cluster (Average Kva = 87) is identified as TechCode is "Railway Engineering" and DeptCode is "Electrical Engineering".

4.5.3 Association Rules

Procedure for mining of Association Rules includes the following seven steps: (1) Converting data type into transaction data; (2) Linking data table; (3) Selecting potential rules; (4) Setting up supports; (5) Determining screening criteria; (6) Screening out significant association rules. After DM, 11 meaningful association rules were obtained. The rules are named A1~A11. An example of the Window of the mined association rules is shown in Figure 5.

Examples of rules obtained from Association Rules are: (1) A3—If ForumID is "Steel Community" and Kshare is "Wisdom", then Kva is >= 80 (Support=20, Confidence=0.5); and (2) A4—If ForumID is "Bridge" and Kashare is "235C00" ("structural" technical domain), then Kva is >= 80 (Support=10, Confidence=0.5); (3) A5—If ForumID is "Geotech" and TechCode is "483I10" ("Viaduct rail" technical domain) and DeptCode is "MRT", then Kva is >= 80 (Support=20, Confidence=0.5);.

4.6 Performance Improvement Strategy Planning

Based on the DM results described above, four strategies for KMS performance improvement are induced as follows:

Strategy I: Establish Screening Criteria for CoP

Strategy descriptions: "In order to eliminate the non-relevant articles, a screening system should be established. In the meanwhile, facility for uploading prevailing news and events should be provided".

Facts: D1, D2, C1, C2, C4, A1, and A2.

This strategy is induced based on observations of that the average Kva of "Hydro-environ" CoP is obviously lower than other CoPs. After reviewing the KM activities of that CoP, a great amount of articles posted on the CoP are related to prevailing news and events that the initiator would like to share with the CoP members. This situation has affected the members' attitude that tends to ignore the articles, and caused malfunction of the CoP. It is therefore recommended to screen out those articles from CoP and to provide the KM initiators (information providers) a facility to upload those articles directly onto the knowledge database of the KMS.

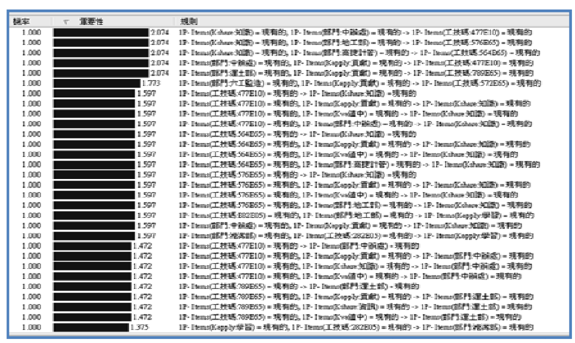


Figure 5 Window of the mined association rules

Strategy II: Establish a New Forum for Rail Related Disciplines

Strategy descriptions: "Railway and Mass Rapid Transportation (MRT) related issues contribute to high Kva values but those issues usually need be solved by multiple disciplines. It is recommended to establish a new CoP consisting of members from architectural, structural, geotechnic, electrical, and project management disciplines."

Facts: D11, C8, C10, and A7.

This strategy is recommended based on the observations that the Kaohsiung MRT project related articles were posted in sevral different CoPs and responded by the members with different backgrounds. Most of the KM activities were evaluated with high Kva. By interviewing with the KM initiators (questioners), it is found that the railway or MRT related engineering issues are highly diversified. Disciplines from different backgrounds are required to incorporate in the problem-solving process. This is actually the Medici Effect as reviewed previously.

Strategy III: Enhance the Lessons Learning Functions for Supervision Departments

Strategy descriptions: "Supervision departments and offices contribute significant amount of high value added KVAP activities. Lessons learning functions should be provided directly to those departments and offices to capture the valuable lessons and experiences from the construction sites."

Facts: D9, C8, C12, and A11.

This strategy is recommended based on the observations that the KVAP activities (Kshare) are evaluated relatively highly value-added in supervision works and issues. By reviewing the KM activities, it is found that those issues were encountered on construction sites and needed to be solved promptly. The results of problem-solving were also instantly fed back from the construction sites. Therefore, lesson-learning processes are actively performed in these departments and offices. However, due to the over-loaded works and shortage of staffs, the lessons-learned were mostly not recorded. This is absolutely a great loss of the firm. Facility and functions should be provided to those departments to enhance the lessons learning process in order to retain the valuable intellectual assets of the organization.

Strategy IV: Enhance Design and Construction Integration

Strategy descriptions: "Design and construction team usually benefit each other in KM activities. Mechanism should be established to enhance the design and construction integration."

Facts: D7, D9, D13, C3, C7, C8, C10, C12, C13, C14, and A5.

This strategy is recommended based on the observations that many high Kva KM activities (mostly with Kva values >= 80) were those initiated by the site engineers and responded by the design engineers, or reversely. It's obviously that integration between design and construction engineers is beneficial to knowledge value adding. Thus, mechanism (e.g., job rotation, seminars, and lessons-learned conference) can be established to enhance integration.

Conclusions

Many construction organizations have adopted Community of Practice (CoP) as an approach to facilitate the process of knowledge generation and sharing. However, most of previous efforts were spent on hardware and software investments. Prior research has developed models of quantitative performance measurement for a CoP, but there have been no effective method for systematic performance improvement of the CoP.

In this paper, a methodology is proposed to mine knowledge from the KM value-adding performance data for planning effective improvement strategies. The proposed methodology comprises of two major elements: (1) a quantitative model for measurement of the performance of KM activities in a CoP; and (2) the commercial DM software—Microsoft SQL Server®— for performing DM tasks. A case study was conducted for nine selected CoPs of a local leading consulting firm in Taiwan. Totally, 4,892 historic KM cases are sampled for questionnaire survey. Finally, 4,825 complete and valid responses were obtained for data mining.

Three DM techniques (Decision trees, Clusters, and Association Rules) were employed for mining of the rules and patterns existing in the performance data. Meaningful rules, useful patterns, and important association rules are found by DM. It is summarized that 13 meaningful rules were obtained from decision trees; 14 useful patterns were identified from clusters; and 11 important association rules were concluded. Four performance improvement strategies are then inferred and planned based on the knowledge discovered from the historic performance data. It is concluded that the proposed method provides the managers of KMS and the firm an effective tool to plan effective improvement strategies for their KM initiatives.

Acknowledgement

The founding of this research project was partially supported by the National Science Council, Taiwan, under project No. NSC 97-2221-E-216 -039. Sincere appreciations are given to the sponsor by the authors. The valuable case study information presented in this paper was provided by CECI, Taipei. The authors would like to express sincere appreciations to the CECI, too.

References

[1] Yu, W. D., and Chang, P. L. (2005) "Performance evaluation of the construction knowledge management system—a case study of an A/E consulting firm," Proceedings of ICCEM 2005, Session 4-D, Oct. 16~19, 2005, Seoul, Korea, 6 pp. 1058-1063.

[2] Yu, W. D., Chang, P.L. and Liu, S. J. (2006a) "A Model of Performance Improvement Strategy Planning for a Construction Knowledge Management System," Proceedings of The Fifth International Conference on Engineering Computational Technology (ECT 2006), Session VII—Modeling of Decision Making and Risk Analysis in Construction Management, Sept. 12~15, Las Palmas de Gran Canaria, Spain, 18 pp.

[3] Yu, W. D., Chang, P.L. and Liu, S.J. (2006b) "Quantifying Benefits of Knowledge Management System: A Case Study of an Engineering Consulting Firm," Proceedings of International Symposium on Automation and Robotics in Construction 2006 (ISARC 2006), Session A4—Planning and Management (1), Oct. 3~5, 2006, Tokyo, Japan, 6 pp.
[4] Cabena, P., Hadjinian, P., Stadler, R., Verhees, J., and Zanasi, A. (1998). Discovering Data Mining: from Concept to Implementation, Prentice-Hall, NJ.
[5] Fayyad, U. and Uthurusamy, R. (1996). "Data mining and knowledge discovery in databases." Commun. ACM, Vol. 39, .pp. 24-27.
[6] Han, J. and Kamber, K. (2000). Data Mining: Concepts and Techniques, Morgan Kaufmann Publishers, San Diego, U.S.A.
[7] Leu, S.-S.; Chen, C.-N.; and Chang, S.-L. (2001) "Data mining for tunnel support stability: neural network approach," Automation in Construction, Vol. 10, No. 4, pp. 429-441.
[8] Mezher, M.; Abdul-Malak, M. A., Ghosn, I.; and Ajam, M. (2005) "Knowledge Management in Mechanical and Industrial Engineering Consulting: A Case Study," Journal Management in Engineering, ASCE, Vol. 21, No. 3, pp. 138-147.
[9] del-Rey-Chamorro, F. M., Roy, R., van Wegen, B., and Steele, A. (2003) "A framework to create key performance indicators for knowledge management solutions," Journal of Knowledge Management, 7(2) 46–62.
[10] Bassion, H. A., Price, A. D. F., and Hassan, T. M. (2005) "Building a conceptual framework for measuring business performance in construction: an empirical evaluation," Construction Management and Economics, 23(5), 495-507.
[11] Yu, W. D., Yao, H. H., Liu, S.J., and Chang, P.L. "Knowledge Value Adding Model for Quantitative Performance Evaluation of the Community of Practice in a Consulting Firm," Proceedings of The Sixth International Conference on Engineering Computational Technology (ECT 2008), Session VIII—Decision Making in Engineering Management, Sept. 2~5, Athens, Greece, 18 pp., 2008.
[12] Mitra, S., Pal, S. K., and Mitra, P. (2002) "Data mining in soft computing framework: A survey", IEEE Transactions on Neural Networks, 13(1), 3-14.
[13] [Furnkranz, J., Petrak, J., and Trappl, R. (1997) "Knowledge discovery in international conflict databases", Applied Artificial Intelligence, vol. 11, 91-118.

A Component-based Approach for Generation of 3 Dimensional CAD Models Using a Construction MODEL Component Database

Hung-Ming Chen[1] and Yi-Chi Chu[2]

[1]Associate Professor, Department of Construction Engineering, National Taiwan University of Science and Technology, 43, Sec.4, Keelung Rd.,Taipei,106, Taiwan, e-mail:hungming@mail.ntust.edu.tw
[2]Graduate Research Assistant, Department of Construction Engineering, National Taiwan University of Science and Technology, 43,Sec.4,Keelung Rd.,Taipei,106,Taiwan, email:M9505502@mail.ntust.edu.tw

Abstract

In recent years, three-dimensional visualization technologies have been researched and applied for aiding the operations of design, planning, construction, etc. in construction engineering. A three-dimensional (3D) CAD model of a targeted facility or construction project is a requirement. Due to the uniqueness and custom features of construction products, their 3D models usually have to be built based on basic geometric modeling elements from scratch. Generally, this is a time-consuming and labor-intensive process. In addition, the reusability of built 3D models is low. For improving efficiency in building 3D models and reusability of built 3D models, this research proposes a component-based approach for generation of 3D models using a construction model component database. A construction model component database is developed to store and provide 3D models of construction components, such as materials and equipments, to be queried and retrieved for generating global 3D models in construction engineering. A management interface of the database is developed for inserting and editing construction component models. A model building interface which utilizes the database for generating global 3D models based on search, selection, and then setting operation is also developed. The data format of 3D component models in the database is X3D for interoperability. Base on the proposed approach, 3D CAD models can be built more efficiently by instancing existing components in the database. In addition, built models can be decomposed into component models to be stored in the database for being reused.

Keywords: database; Visualization; 3D model; CAD; Virtual Construction.

1. Introduction

Recent years have seen the introduction of 3D CAD to the design, planning and execution phases of civil engineering projects. In the design phase, constructing a 3D model allows planners to visualize the entire project, and helps them to quickly find and resolve design conflicts. A 3D model can also help ensure that complex procedures fully conform to stated design principles. If each stage of construction is represented in turn, the three-dimensional model becomes a four-dimensional animation of the construction process. This animation can give the engineer the experience of seeing the construction process before it ever happens. McKinney and Fischer (1998) suggest that this process can overcome the problems of traditional, two-dimensional plans. Clayton et al. (2002) also think that virtual sites based on 3D models can help understanding of the engineering situation.

3D models can also be extended for virtual testing and performance prediction. They can be used to test flood prevention systems in reservoirs, or predict building structure performance during earthquakes. For example, Chen and Hsieh (2008) propose constructing 3D models complete with information on materials, stress, etc., then using XML transformations to integrate them with commercial analysis software. These applications are a unique function of 3D CAD, and cannot be replicated using traditional 2D designs.

However, the products used in construction are highly specialized and customized, so 3D modeling has always been a labor- and time-intensive process. Once developed, a model generally cannot be used on more than one project. 3D modeling is also expensive, so despite its advantages, it is not yet a standard procedure in the construction industry. This paper proposes establishing a database of reusable components for 3D models, which would allow for modular construction. This database would cut modeling time, costs and labor.

2. Research Objective

Buildings and bridges, the products created by the construction industry, are generally unique. They are designed to meet specific needs, and the constraints that the site environment places on each construction will be different. Therefore it is current practice in the industry to develop each 3D model from scratch. Appropriate software is selected based on the design requirements, and the specific features of the design are input by engineers. There is currently no widely applicable mechanism for reducing the investment of time and labor in setting up a 3D model. Once developed, the majority of models cannot be re-used.

If a model could be decomposed into components, and those components could be reused in the development of other models, it would save a lot of time that is currently wasted on building similar models anew each time. This paper proposes two innovations: (a) a database of 3D model components which can be put together in a modular fashion; and (b) a development process that uses these components to construct full 3D models. The development process involves three phases. (1) The user searches the database for components that meet the project requirements. (2) The desired components are selected and pieced together in a 3D environment. (3) The properties of the components are adjusted as necessary. The steps are repeated as often as necessary until the selected components fit together as a single model that meets the project design requirements. The process is illustrated in Figure 1.

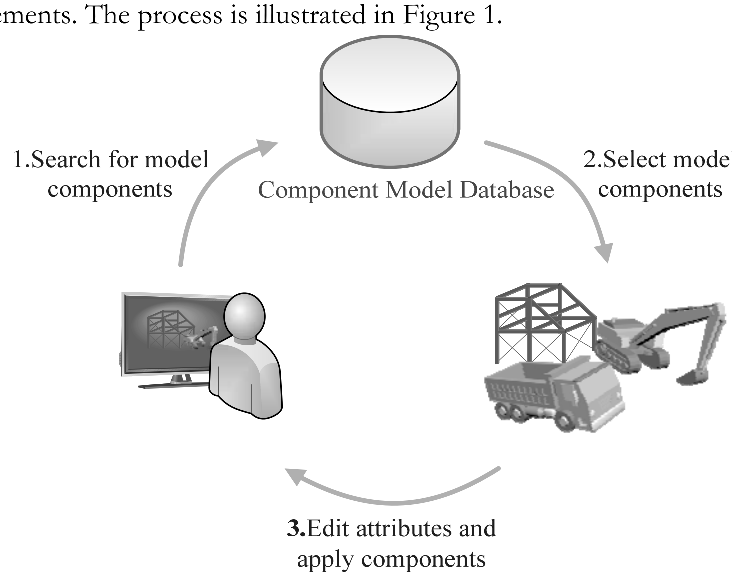


Figure 1. Model development process

3. Format

The components are of three basic types: personnel, equipment systems, and materials. They are defined by their geometric properties.

It is desirable that the format for components should be an open standard, not a proprietary format linked to a certain software. If the database can be searched and added to from any software platform, it will have greater potential for fast development. Therefore, the royalty-free, open X3D standard was chosen as the format. X3D is a new XML-based format developed out of VRML. It features a rich system of tags, is highly extensible, and is very compatible with web applications. In future most 3D design software will support X3D.

4. System Elements

The system needs to provide a database of components; a database management module; model construction module; and a format converter.

4.1. Database Management Module

User input will be necessary if the database is to supply an accurate, varied and complete range of components. Users will add to and enrich the components stored, increasing the flexibility and utility of the system. The system must therefore include a mechanism for user management and maintenance of the database. Necessary maintenance operations include adding new components, deleting components, and editing component features. Users must also be able to manage component classes. The system management mechanisms and functions are illustrated in Figure 2.

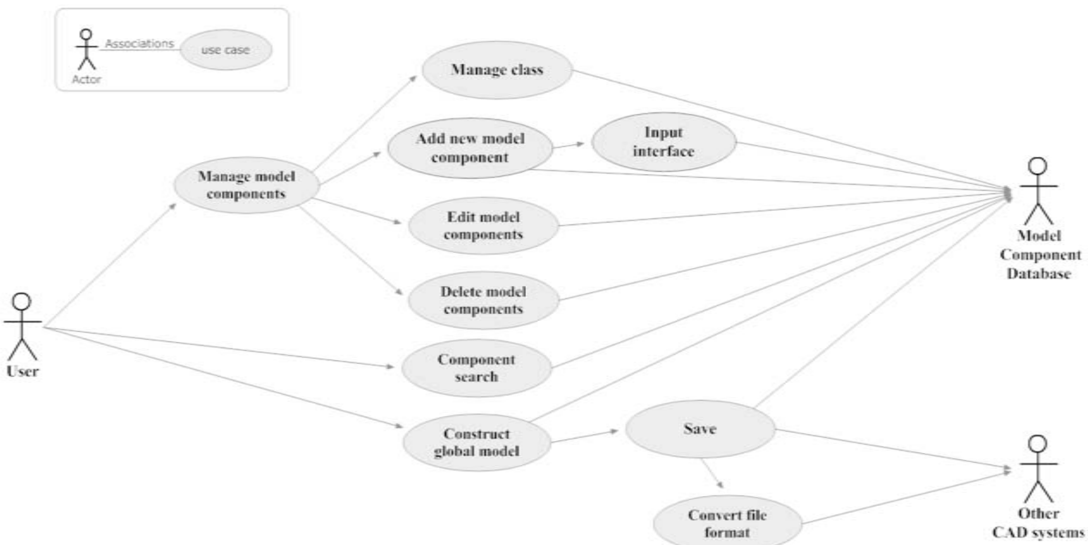


Figure 2. Model component database management system

There will be three ways in which a new component can be entered into the system. (1) Components can be input using the component management module. (2) If a component is already in X3D, the user import it directly. (3) If a component is in an incompatible format, it can be converted through by the format conversion module, and then imported. (See Figure 3).

4.2. Model Construction Module

The system needs a design interface through which users can assemble components into a final structure. This interface must include user-friendly 3D movement and perspective control, and flexible settings for the 3D environment. Most importantly, it must be easy to find suitable components and to apply them in the environment. The search function must be effective enough that there can be a seamless transfer of the user's design concept into the model. It is the combination of the design interface and the component database that will allow for a dramatic rise in efficiency in 3D modeling.

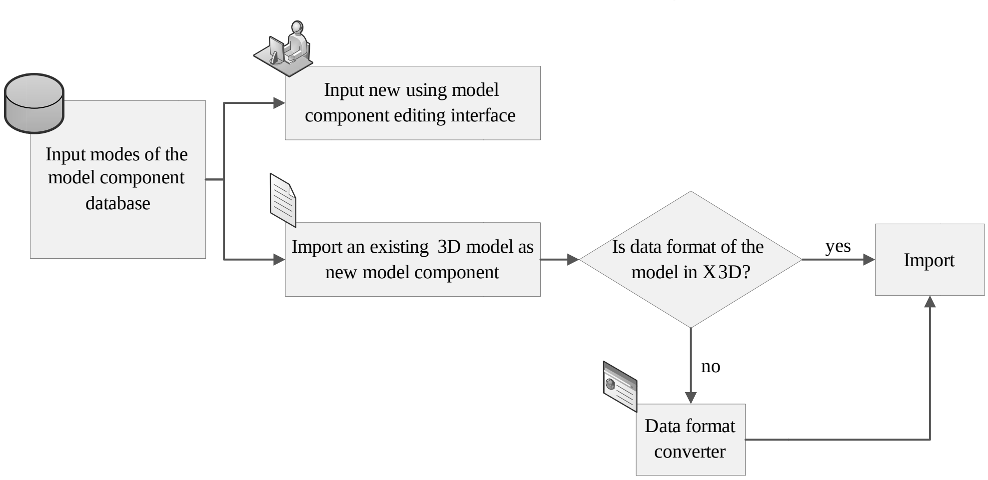


Figure 3. Adding model components to the database

4.3. Format Converter

The decision to base the database on an open standard means that the system will be interoperable with other CAD software systems. The X3D format is based on XML, and it is therefore possible to use XML to automatically convert to other formats. For the purposes of this study, the format converter will be tested by converting X3D components to VRML.

5. Development of the System

The structure of the system is shown in Figure 4. Its major components are the component database and three interfaces: the database management module; the model construction module; and the format converter. The database is online, so that all users can access it through the Internet. Users can also add to the components in the database, making it more powerful and useful. Because it is an open platform, any software will be able to access the database. Even a web browser equipped with the necessary plug-ins would also be able to view components.

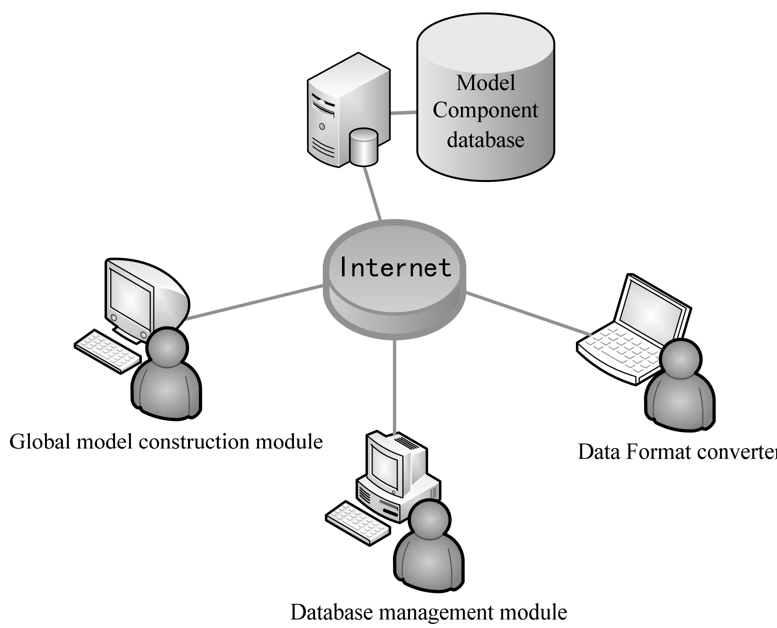


Figure 4. System architecture

5.1. Database Design

Database design must satisfy user demands; for this system, there must be powerful search functions to allow users to quickly find the components they need. Therefore for each component, several basic properties must be entered: component name, author, brief description, small image. Each component will also belong to one or more user-defined classes. For example, a user can define a jib crane as belonging to the class crane, construction machinery, large machinery, etc. Multiple classes allow users to define components in more than one mode, making them easier for other users to find. Users can also define extra properties for components, such as component dimensions. Figure 5 shows the information structure of components in the database.

5.2. X3D DOM Design

X3D documents are composed of nodes and fields. Nodes describe the form and function of a model, fields record the properties of nodes.

In order to allow for direct access to information in the X3D format, the system uses the X3D document object model (DOM). It is object-oriented in terms of accessing X3D documents, meaning the system can directly access data in the documents in runtime.

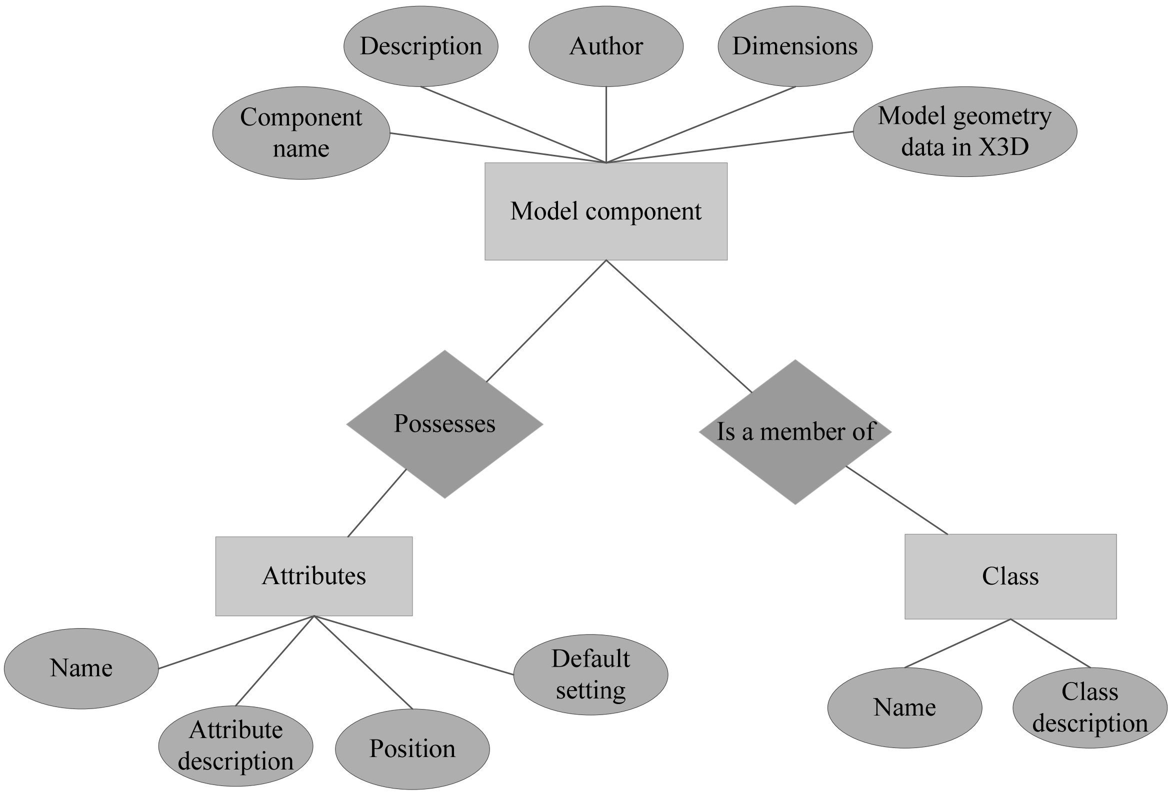


Figure 5. Database schemas

5.3 System Interfaces

5.3.1 Component and class management

The system presents all components in list form, and users can also select the components they want to manage by class. Selecting a component will bring up all of the information for that component (name, description, author, etc.) and a preview. At this point, users can edit or delete the component.

Users can also select Manage classes. The system will then list all classes, and which components belong to them. The user can choose to edit or delete a class, and can also create new classes.

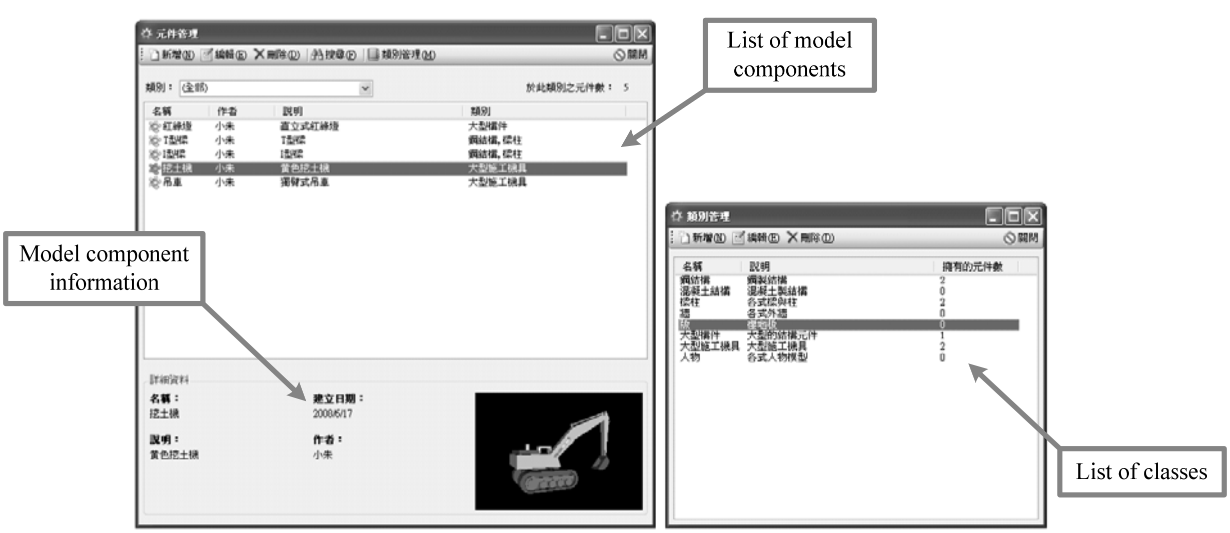


Figure 6. Interfaces for managing model component s and classes

5.3.2 Adding, deleting and editing components

When a user adds a component, they enter the X3D code in the component window. The preview window shows the component as it is entered. The system can recognize the X3D format, and indicates when incorrect code is entered. Users can also select Import X3D file, and import any X3D file as a component.

The X3D community has suggested some conversion tools for other common 3D formats (e.g. 3DMax). 3D models in other formats can thus be converted into X3D and imported.

When the X3D code has been entered for a component, a description and basic properties must be entered for the reference of other users. Properties will also make the component searchable. Basic properties include the name of the component, a description, the name of the author, etc. The component must also be assigned to one or more classes. When this information has been entered, the new component is formally accepted into the system.

Users can also edit existing components by selecting Edit. The edit interface is identical to the component input interface.

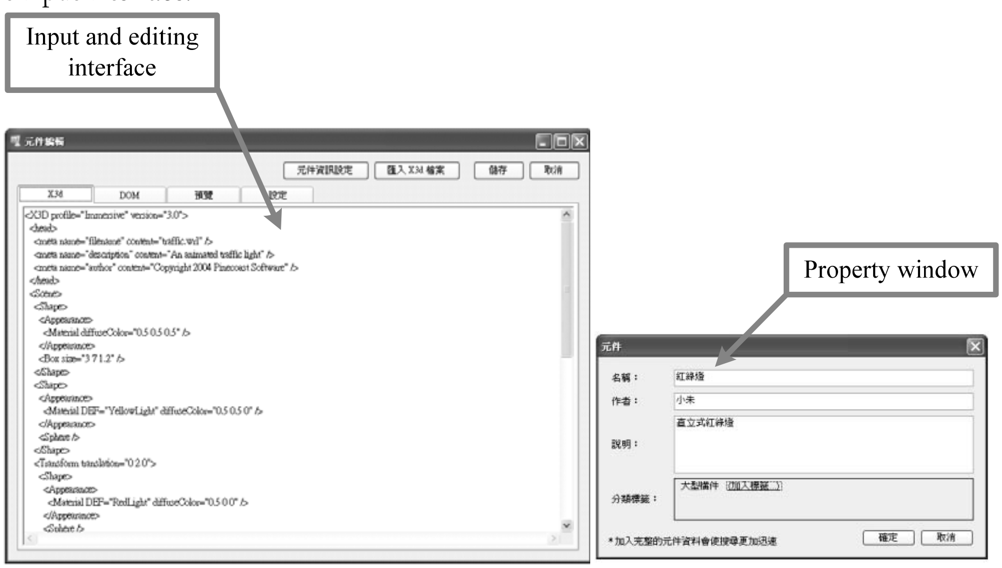


Figure 7. Interfaces for inputting and editing model components

5.3.3 Search

Users need to find components quickly and efficiently when constructing a model. The system allows searches in three modes. (1) Keyword search. The system goes through the database component by component, and all components with information that match the search keywords are listed. For example, if the user searches component names for the search term "wall", then components named "concrete wall", "brick wall", "boundary wall" will all appear in the search results. (2) Class search. Users can search for all components within one or more classes, without entering search terms. (3) Complex search. Users can set multiple search conditions, using both classes and keywords. For example, it would be possible to search for components by author "John Smith" in the classes "construction" and "architectural components".

These search modes should allow users to find the components they need quickly and efficiently. The list of search results includes a preview of the components, to help users decide which components to select.

5.3.4 Model construction

The process of developing a full 3D model is one of continually setting search parameters, selecting components (by double clicking them) and adding them to the environment. In the environment, users simply click and drag the components into position. They can also use the properties window to alter the properties of the components (size, orientation, etc.). When properties have been set, multiple copies of the altered component can be made, without having to set properties for each copy. Users can also alter the environment as necessary, setting background, lighting, ground conditions, size, etc.

When sufficient components have been added and the model is complete, the user can choose to save it as a separate X3D file, or can save it back to the database as a new component. After saving, the user can continue to edit the file.

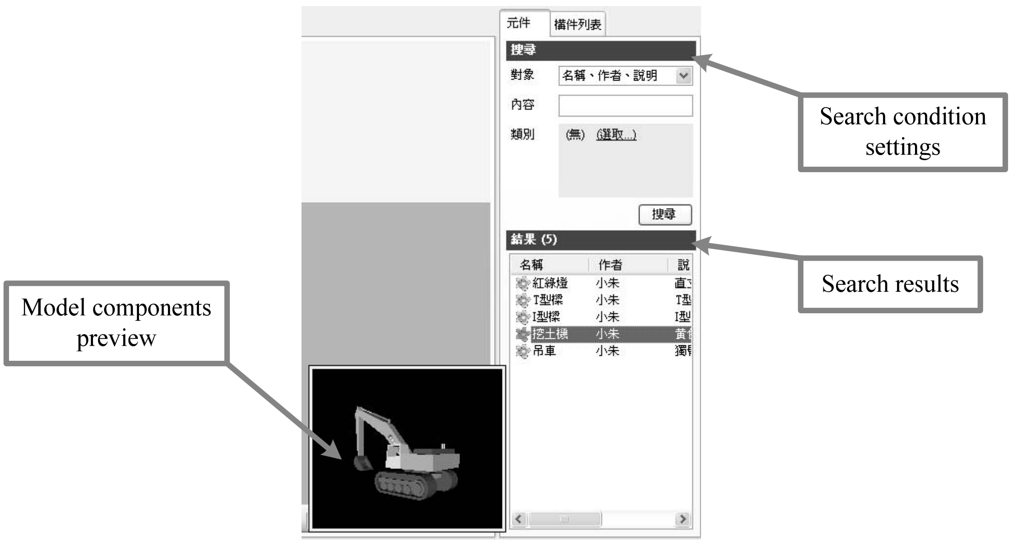


Figure 8. Search interface

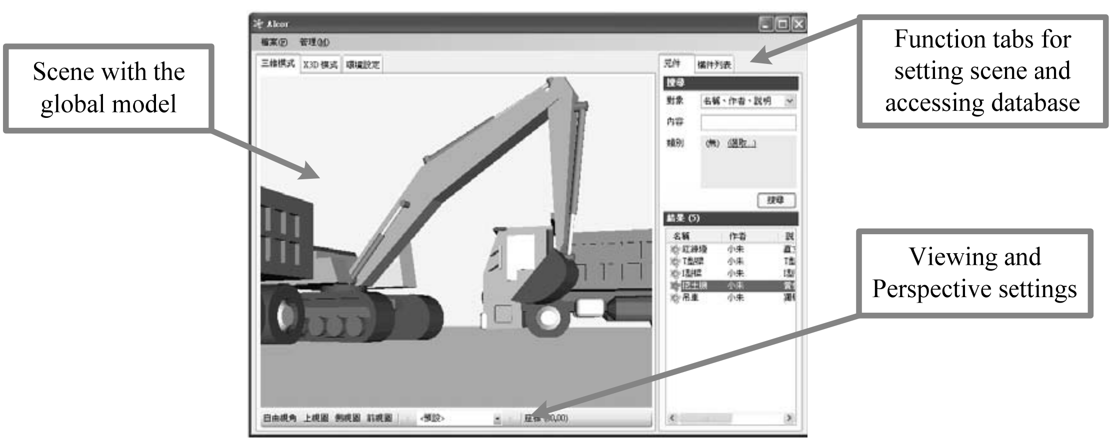


Figure 9. Model construction interface

6. Conclusion

In this paper a new system is proposed for developing three-dimensional models for engineering projects. The system is based on a database of 3D components, with separate modules for model construction and for database maintenance. Using this system, users will be able to construct 3D models more quickly and efficiently. Completed models can be viewed and edited on any software which supports the open X3D standard.

Users also have the option of adding to the component database, allowing future developers to reuse their components. Using preexisting components can reduce the time spent creating a model, and thus increases efficient use of human resources.

There are two major differences between the system proposed here and those already available. The first is the database of 3D components. Using powerful search tools, users are able to quickly and accurately locate suitable components, and apply them in a 3D model. Users also have the option of editing and adding to the database, increasing its power and flexibility. Second, the system uses the X3D open standard rather than a proprietary standard. Most systems or software will support X3D, and it can be converted automatically for those that do not. This means that the database will be accessible to users of almost any system.

References

[1] Chen, H.-M., and Hsieh, J.-W. (2008). "Integration of Design and Analysis Systems Using XML Transformation." Proceedings of the 12th International Conference on Computing in Civil and Building Engineering, Beijing, China, October 16-18, in CD-ROM.

[2] Clayton, M.J., Warden, R.B., and Parker, T.W. (2002). "Virtual construction of architecture using 3D CAD and simulation." Automation in Construction, 11(2), 227-235.

[3] McKinney, K., Kim, J., and Fischer, M. (1998). "Visualization of construction planning information." Proceedings of the 3rd international conference on intelligent user interfaces, January 6 - 9, 135-138.

A 3-dimensional Visualized Approach for Maintenance and Management of Facilities

Hung-Ming Chen[1] and Yu-Hsiang Wang[2]

[1]Associate Professor, Department of Construction Engineering, National Taiwan University of Science and Technology, 43, Sec.4, Keelung Rd.,Taipei,106, Taiwan, e-mail:hungming@mail.ntust.edu.tw
[2]Graduate Research Assistant, Department of Construction Engineering, National Taiwan University of Science and Technology, 43, Sec.4, Keelung Rd., Taipei, 106, Taiwan

Abstract

The maintenance and management of facilities is an emerging issue. Decisions on maintenance-related works usually are made based on various types of accumulated historical data, such as design drawings, inspection records, sensing data, etc. There are systems for storing and maintaining such maintenance-related data electronically in database. However, the data accessing mode of these systems is based on text input on web form, which sometimes is not intuitive enough for interpreting retrieved information for decision making. In addition, most systems are only developed for one type of data, and for a specific facility. This limited the completeness and extensibility of these systems. This research proposes a 3-dimensional (3D) visualized approach for maintenance and management of facilities. A prototype system which can apply the proposed method to different facilities is developed for concept proofing. A 3D facility model is provided in the system as the interface for accessing various maintenance-related data intuitively. In addition, the presentations of various maintenance-related data are visualized on the model as possible to provide user an intuitive understanding about the states of the facility in many aspects. Behind the 3D visualized interface is a database which systematically integrates and stores various maintenance-related data together. The data of this database should be constantly accumulated via input from users and sensors in appropriate and suitable formats, and can be analyzed and handled by available methods, such as reliability and knowledge management, to provide processed information for aiding decision making in maintenance.

Keywords: Visualization; Maintenance; Management; Facility; 3D Model.

1. Introduction

Facility maintenance and management (FMM) becomes relevant once the facility is completed. Proper FMM can help administrators identify problems as early as possible and maintain the facility effectively. FMM normally needs a variety of data for decision making. Therefore, in addition of effective storage of original design data, the operator needs to record the facility's condition regularly, and inspect and keep records of the facility at suitable times. Effective FMM needs to integrate and manage information, such as temperature, inspection records, maintenance records, drawings, and contracts. This provides administrators with enough information to make an FMM decision.

However, FMM data are mostly kept by handwritten record books or repair records. Administrators need to check paper records and calculate the maintenance schedule manually. Under this FMM mode, records are kept "on paper" and are not digitized. In addition, the current FMM is divided into different aspects. Maintenance records are also possibly distributed to different places. For example, maintenance information and inspection records are kept and managed separately and do not link with each other even if they are highly associated in making decisions to maintain a specific facility. Discrete information is not integrated, and administrators cannot make the optimum FMM decisions based on relevant FMM data. Looking through massive volumes of maintenance records is labor- and time- consuming. In time-sensitive or emergency situations, it would be difficult to provide the latest FMM information or integrate it effectively.

Therefore, some researchers introduced information technology (IT) into facility maintenance mode, digitized maintenance and management (MM) information, and even integrated related data, which is convenient for administrators. Examples of such integration in Taiwan are the Facility Graph Maintenance and Management System developed in Sinotech Engineering Consultants, Ltd (2006), the Taiwan Bridge Maintenance and Management System developed for the Ministry of Transportation in Huang (1997) and Lee (2005). These systems provide digitized information for inquiry, and present results in forms. However, this presentation mode is not intuitive for some types of data, such as facility temperature distribution and facility location. When administrators read digitized data or information, they cannot immediately get the distribution and state of the entire facility through the mode, and can only understand it by virtue of experience. Additionally, existing systems put emphasis on only one aspect of maintenance information, for example, the system in Sinotech Engineering Consultants, Ltd (2006) only carries out MM for facility graphs and textual information, and the system in Huang (1997) and Lee (2005) only manages the maintenance and inspection records for bridge facility. In those systems, other information such as contractual documents; monitoring information is not integrated.

As for the subsequent application of the three-dimensional (3D) model into FMM, although some applications have already applied 3D visual technology to visualize the facilities targeted for MM, such as the spatial navigation systems established for museums, campuses, and cities as in Tamada et al (1994) and GSPRS (2005). However, the functionality of visualization in these systems is limited to navigating the overview of the facility. In addition, they have not been applied or incorporated into the operation of FMM, and have no linkage to the maintenance information of target facility.

Apart from this, the design and establishment of most existing MM modes are customized to a specific facility and cannot be applied to others. This wastes resources and time. As for the reusability of maintenance information, most existing maintenance modes are used to record maintenance information alone. In practice, some maintenance data can be further processed to produce useful information for FMM. Recently, some researchers such as Dell'Orco & Zambetta (2003), Billinton & Abdulwhab (2001), Donaghy & Omanson (1989), and Lee (2007) conducted studies on FMM based on post-processing of information, but did not adopt visual technology or integrate other MM information.

To sum up, this study proposes a new 3D visual FMM approach to expand the current MM mode. The new approach takes advantage of visual technology to provide effective search and presentation of maintenance data, uses a single database to integrate and link a variety of FMM information, and effectively uses knowledge management, artificial intelligence, and other technologies to analyze and process information to support decision making. All of these allow the new mode to become a superb FMM mode and system featuring 3D visualization, integration of data and monitoring information, information analysis, and decision-making function.

2. Research Objective

The objective of this study is to propose a 3D visual FMM approach, and its concept is illustrated in Figure 1. This approach gives the facility a 3D visualized view, which allows administrators to view the virtual facility and the scene, as well as press buttons to rotate, move, or zoom the object in and out. This way, administrators can easily view and click the part needed. Administrators can also easily select each component of the facility, and obtain the maintenance or management information required. The selected part will change color and display the basic information. In addition, the MM information for each facility or component is available, data are presented in the most suitable way, and both single selection and multi-selection are supported.

This approach stores FMM information in a proper digitized format such as image, data, or file, and links information to the 3D model of the facility. When administrators click "Facility Inquiry," all detailed basic data and historical records of this facility will be provided. In addition, this approach has integrated diversified MM information sets such as basic information, examination records, monitoring information, and others. It uses a single database to store and manage information collectively, and presents information through a 3D visual interface so that the monitoring information will directly reflect the real-time state of the facility by color or action, such as facility temperature and lift location. Furthermore, the post-processing of MM information is carried out with knowledge management, artificial intelligence, and other data processing technologies. For example, based on maintenance and measurement information, we can computerize facility failure rate and reliability through statistical formula, and then make a maintenance plan for the facility. Through the knowledge management mode, this mode can sort and present MM information properly, and the result of processed information can effectively bolster decisions on maintenance work.

(a) 3D visual interface

(b) Design and maintenance history of digitized storage facility

(c) Integrated MM and data collection and storage

(d) MM information processing and analysis

Figure 1. Conceptual illustration of the FMM approach proposed by this study

3. System Requirement Analysis

This research focuses on the system development for the aforementioned research objective, carries out a study on functional requirements, and develops a system plan. For the system's functional requirements, we analyzed and planned from the six aspects of data type, data resource and digitized format, database skeleton, data presentation, treatment of data hierarchy, and the building mode of facility management 3D model. The following sections give an overview of the analysis and planning based on six aspects.

3.1. Data Type

By understanding and analyzing FMM operation, we decided to use the system to store and manage this type and scope of facility data, as shown in Table 1. The data can be divided into six categories according to their nature, and the coverage of each type is listed in Table 1. When maintaining and managing the facility, administrators need to inquire about the basic information on the facility to determine its situation. From here, they can view the engineering drawings to know the facility's structural design and engineering plan. Administrators will also periodically examine the facility and keep records for subsequent follow-up. Records for damaged facility after repair should also be kept and stored for later review. Combining this with a facility monitoring device will capture real-time, rapid, and dynamic information about the facility, which is beneficial for keeping track of its latest state. FMM work is also expected to optimize efficiency, reduce MM cost, and make MM work stable and reliable. However, the scope of the six kinds of MM information is very large, and this study will establish the prototype system for a few items of maintenance information for each kind to explain the concept proposed by this study.

3.2. Data Resource and Digitized Format

Based on the type of data to store, this study analyzed the data resources and determined the suitable digitized format for future planning of a suitable system interface, data collection, and data storage and management in the database, as shown in Figure 2. It is divided into four parts. The system requires a network client input interface, which uses manual filling or uploading to enter facility information, facility thumbnail, examination records, maintenance records, and incident experience into the database individually for future inquiry. Facility drawings and contractual documents will be uploaded into the system and recorded in the database as scanned files. As for the real-time state of the facility received from monitoring equipment, such as room temperature decreases due to air conditioners, lift locations, and facility fire alarm, they will be recorded into the database and will be presented through the 3D interface. Finally, a statistical analysis of maintenance records such as a statistical comparison of maintenance frequency will compare each facility. Facility failure rate, reliability, and other information will also be obtained through statistical or historical data. All analyzed information will be stored into the database for further MM.

Table 1. Data type and example

Data type	Example
Basic information of facility	Facility name, builder, designer, power consumption, region, and facility category
Engineering drawing of facility	Elevation drawing, sectional drawing, plan view, perspective drawing, construction drawing, and electromechanical equipment drawing
Examination record	Equipment records, operation state, daily workload, and time cycle
Maintenance record	Maintenance and accident records, failure mode, and maintenance record
Facility real-time monitoring information	Temperature, humidity, stress, strain, displacement, elevator position, and fire alarm
Analytical result	Equipment reliability, optimum maintenance schedule, and system failure rate graph

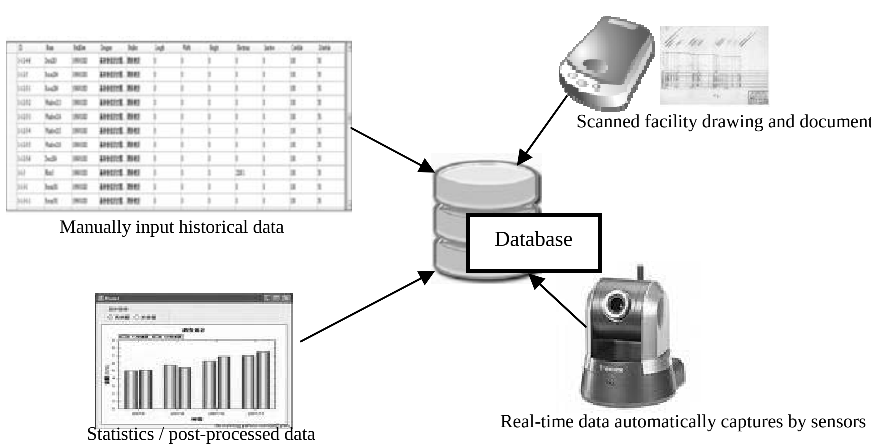


Figure 2. Data sources and digitized format

3.3. Database Schema

This system needs to incorporate all kinds of available facility maintenance information into one database, and obtain the overall integrated information of the facility through a 3D visual interface. This study is based on the correlation between the data required to store and the digitized format. It planned and designed the database skeleton. The information is stored into several datasheets according to type and the facility's basic information. This study has also established the data's dependency according to their correlations, as shown in Figure 3. It takes the facility object as the core and associates it with relevant incidents, maintenance record, monitoring information, empirical files, and the facility's basic information. By using such design, administrators can access all information or states of the facility in the database by clicking the facility on the 3D interface.

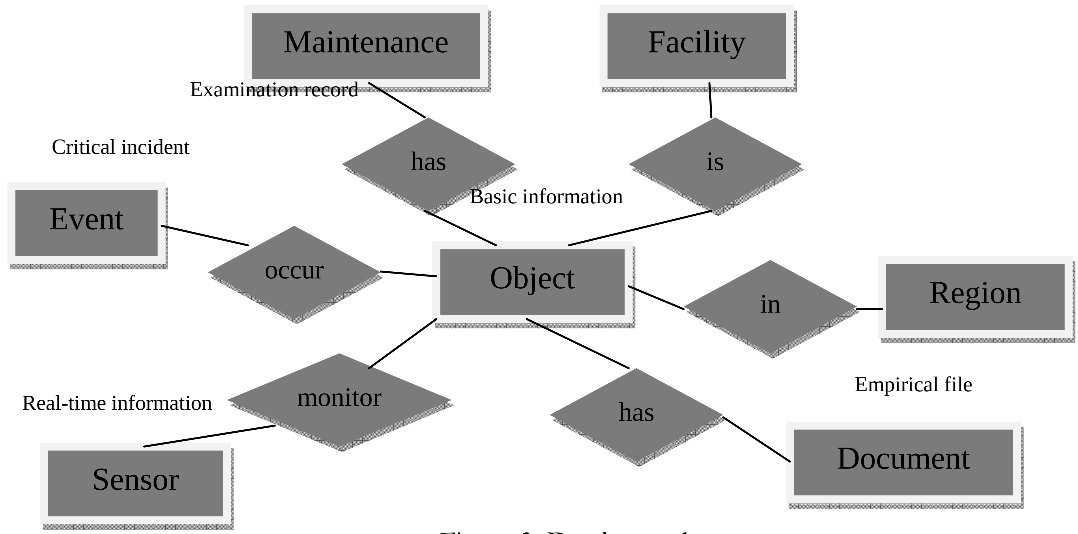


Figure 3. Database schema

3.4. Data Presentation

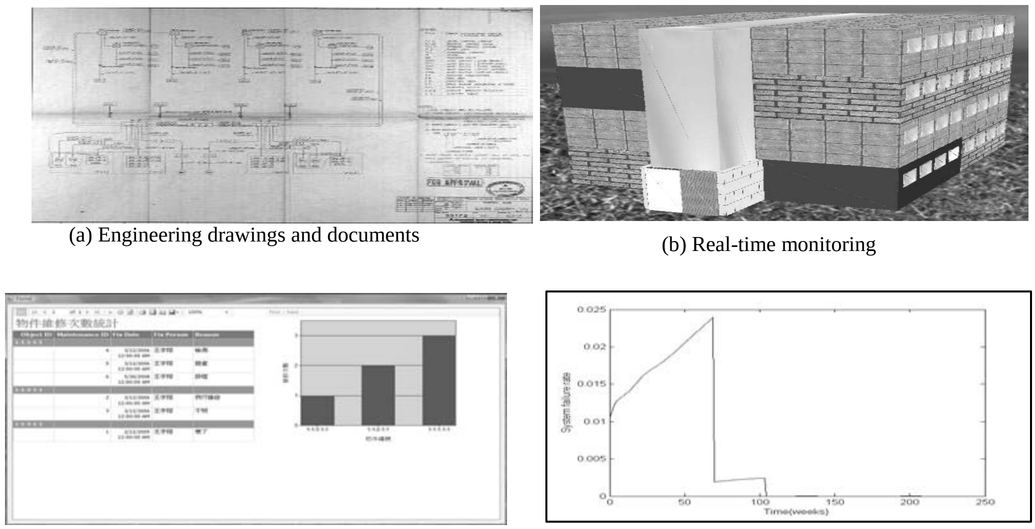
FMM information has different types and aspects, and each kind of information has its own suitable presentation in view of its unique category. According to the information of different types and aspects, this study proposes the best visual presentation mode, and allows administrators to understand it effectively and intuitively and make decisions by considering comprehensive facility maintenance information. The visual presentation modes proposed by this study for the maintenance data of each kind of facility are shown in Figure 4.

Most engineering drawings and historical information are text or graphs, and conventional management is used to manage them through text or form. Therefore, this study takes "form and text inquiry" as the presentation of such information.

Real-time monitoring information is mostly monitoring data, such as equipment temperature, lift location, and incident occurrence received from the monitoring equipment. If they are presented simply in text or form, administrators can only access the data, which is not very intuitive. Since this study supplies a 3D interface, the monitoring information is also presented by the 3D interface to allow administrators to get close to the actual situation of the equipment. For example, for facility temperature in a 3D interface, it will use the distribution of colors to present the current temperature of the facility. This way, administrators can access the temperature of the facility and compare it with adjacent equipment to identify abnormal temperature. As for lift location, real-time movement is seen in the 3D interface and allows faster follow-up in the floors where the lift is located. In the event of a fire alarm, the 3D interface will warn administrators by blinking. The statistical information is the statistical result of maintenance record or historical information, and the best presentation of the statistical information is by statistical statement. Therefore, this study adopts a statistical table to present such information. Post-processed information refers to information obtained after processing through knowledge management or artificial intelligence mode. In our study, we find that some information obtained is suitable to present in 3D interface, while some are suitable to simply present in text. For example, the analytical result of failure rate is presented in the 3D visual interface through distribution of color; warning of the maintenance schedule is presented in 3D visual interface.

3.5. Processing of Data Hierarchy

The match between FMM data and its model has a hierarchy problem. The facility consists of components in different hierarchical dimensions. For example, a building can be divided into floors, and a floor can be subdivided into rooms. The facilities in different dimensions correspond to respective MM information, that is, the building has information on power consumption; the room has information on temperature; and the lift has information on location, which the building does not have. Therefore, this study develops a solution for data hierarchy and allows administrators to switch between and select the components of facilities in different dimensions to inquire about related information. This study adopts the identification (ID) design for the facility model component to solve the data hierarchy problem. Each facility or each component in different dimensions in the facility has a unique ID, with a format of Project No.-Building No.-Floor No.-Room No.-Object No. (-Object Part No.).



(a) Engineering drawings and documents

(b) Real-time monitoring

(c) Statistical information

(d) Post-processed information

Figure 4. Presentation of data

Figure 5. Building Steps of the Model

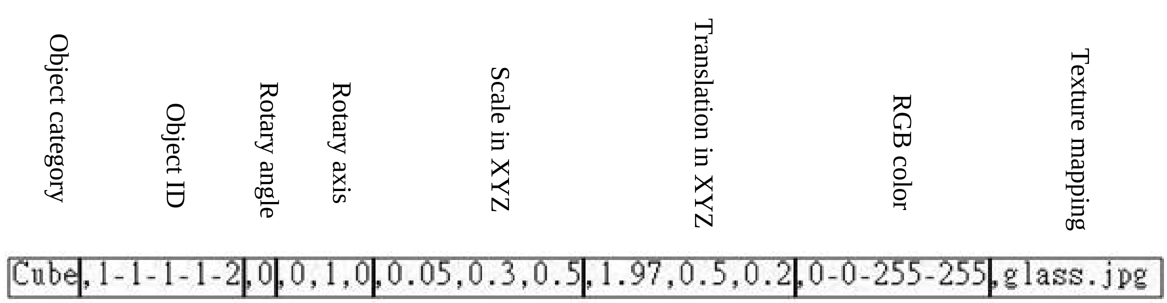


Figure 6. File format and instance

3.6. Construction Mode for the Facility Management 3D Model

The 3D facility management mode proposed in this study is created by using the prototype system developed by this study based on OpenGL and VB.NET 2.0 (see Figure 5 for the establishment of the facility management 3D model and mode in relation to varied maintenance information in the database).First, the model builder needs to assign a unique ID to the facility model or its component, (ID is the hierarchical code of different dimensions), then it uses OpenGL libraries to draw the 3D model based on the scale and material of the facility and finally links the ID of these model components to the information recorded in the database. This way, the creation of the facility model is completed. Additionally, in this study, we developed a file format to allow users to store the models built for different facilities, as shown in Figure 6.

Lastly, this study uses OpenGL Select functions to realize single and multi-selection of objects. If objects overlap, we select the upper object and apply rotary or deformed view to select the covered object.

4. System Framework

The 3D FMM system platform proposed by this study integrates the four technologies of monitoring, visual interface, 3D model, and data analysis and processing, as shown in Figure 7. The 3D facility model is established based on the application of the concept of Building Information Modeling (BIM) to facility MM. BIM contains attribute information, geometrical information, and correlation of the 3D model; the presentation of visual interface adopts the 3D drawing technique. It draws the visual interface according to BIM information and presents it through animation. The monitoring takes advantage of the data received from a dynamic access monitoring equipment, links to information in the FMM database, and uses the visual interface to present. Finally, it uses knowledge management and artificial intelligence techniques to analyze and process the data. The result can support decision making within the MM. It is estimated that this integration will optimize MM decision making and simplify the experience.

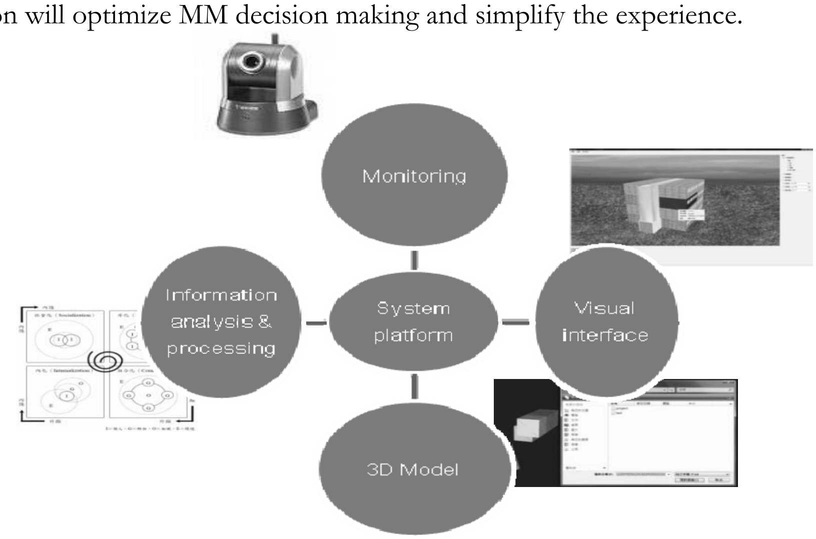


Figure 7. Technologies integrated in the proposed approach

The structure of the prototype system developed by this study is shown in Figure 8. The entire system platform is created based on four aspects. Administrators use 3D visual interface to accomplish facility MM. By inquiring in the back-end database, the system presents information for administrators to view. The monitoring part transmits the monitoring information through the network. The facility model draws established facility items by OpenGL visual technology to allow administrators to click.

5. Conclusions

This study applied and integrated visualization and database technologies into FMM work, and proposed a new FMM mode featuring (1) Integrity, capacity to integrate MM information, (2) intuition, allowing administrators to carry out FMM intuitively, (3) real-time dynamics, observing the current state of the facility by facility monitoring, and (4) reusability in which this mode can be used repeatedly via the facility model building. This 3D visual FMM mode can expand the current FMM mode by addressing inadequate digitization, difficult integration of relevant data, and unintuitive data presentation. The system developed by this study can be used for general projects, and the aforementioned model can be effectively applied to other facility management projects repeatedly.

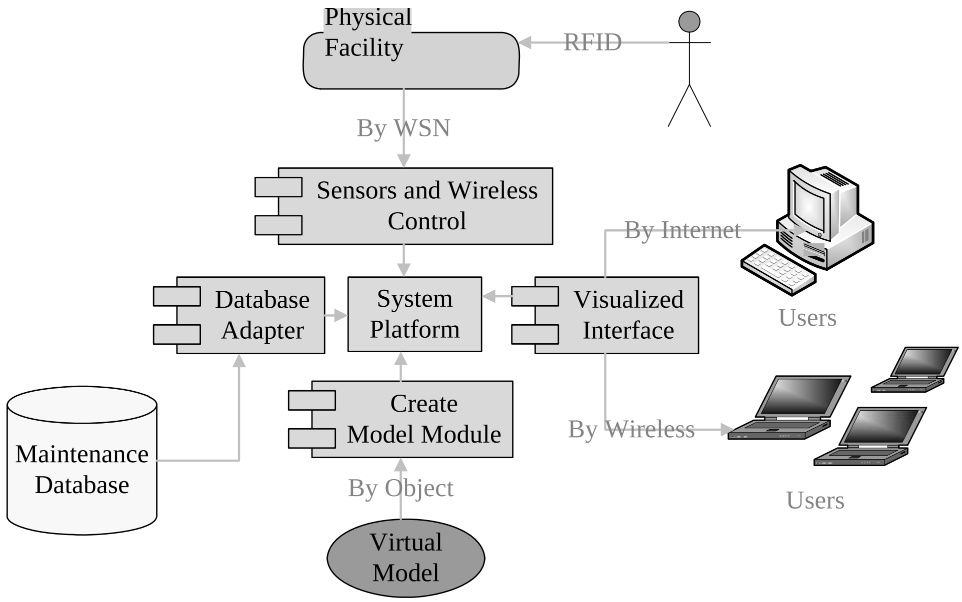


Figure 8. System architecture

References

[1] Billinton, R., and Abdulwhab, A. S. (2001). “Application of probabilistic health analysis in generating facilities maintenance scheduling.” Proceedings of the 2001 IEEE Proto Power Tech Conference, Porto, Portugal, September 10-13, 2(6).

[2] Dell’Orco, M., and Zambetta, M. (2003). “An artificial intelligence approach for parking facilities management” Proceedings of the 4th International Symposium on Uncertainty Modeling and Analysis, September 24, 2-8.

[3] Donaghy, J. P., and Omanson, R. C. (1989). “MICE: A facility maintenance expert system” Proceedings of the 1989 IEEE International Conference on World Prosperity Through Communications, June 11-14, 3, 1418-1422.

[4] German Society for Photogrammetry and Remote Sensing, European Spatial Data Research, and University of Bonn (2005). “Next generation 3D city models” Proceedings of the First International Workshop on Next Generation 3D City Models, Bonn, Germany, June 21-22.

[5] Huang, X.-J. (1997). Study on System Architecture for Road and Bridge Management in Taiwan, Master Thesis, National Central University, Taiwan. (in Chinese)

[6] Lee, B.-X. (2007). Study on Application of Data Warehouse in Creation of Building and Equipment Maintenance and Performance Evaluation System, Master Thesis, National Chiao Tung University, Taiwan. (in Chinese)

[7] Sinotech Engineering Consultants, Ltd. (2006). Manual for Document Maintenance and Management System for Facility. (in Chinese)

[8] Tamada, T., Nakamura,Y., and Takeda, S. (1994). “An efficient 3D object management and interactive walkthrough for the 3D facility management system” Proceedings of the 20th International Conference on Industrial Electronics, Control and Instrumentation, September 5-9, 3, 1937-1941.

[9] Yang, Z.-H. (2005). Study on Current Situation and Future Developing Strategy for Bridge Maintenance and Management in Taiwan, Master Thesis, National Central University, Taiwan. (in Chinese)

Workflow Re-engineering for Implementing a 4D Construction Management Tool in a Design-Build Project

Meng-Han Tsai[1], Abdul Matin Md[2], Shih-Chung Kang[3], Shang-Hsien Hsieh[4]

[1]PhD student, Computer-Aided Engineering Group, Department of Civil Engineering, National Taiwan University, Taipei, Taiwan; PH (886) 2-3366-4342; FAX (886) 2-2368-8213; e-mail: d9521002@ntu.edu.tw
[2]Civil Engineer and Research assistant, Computer-Aided Engineering Group, Department of Civil Engineering, National Taiwan University, Taipei, Taiwan; PH (886) 2-3366-4342; FAX (886) 2-2368-8213; e-mail:matin.pwd@gmail.com
[3]Assistant Professor, Computer-Aided Engineering Group, Department of Civil Engineering, National Taiwan University, Taipei, Taiwan; PH (886) 2-3366-4346; FAX (886) 2-2368-8213; e-mail: kang@ntu.edu.tw
[4]Professor,Computer-Aided Engineering Group, Department of Civil Engineering, National Taiwan University, Taipei, Taiwan; PH (886) 2-3366-4313; FAX (886) 2-2368-8213; e-mail: shhsieh@ntu.edu.tw

Abstract

The use of 4D management tool in construction industry has been growing rapidly in recent years in response to soaring demands on efficient construction management. Successful implementation of this tool in an engineering firm for use in real projects is still challenging. This paper shares the experience on implementing an in-house 4D management tool in a large EPC (Engineering, Procurement and Construction) firm that has extended reputation on design-build projects. A new workflow was proposed that takes full advantage of the 4D tool to enhance construction management performance and minimizes impacts on the firm's existing workflow. The workflow development for the firm to implement 4D management tool in real project was done by interviewing key personnel in multiple departments of the firm and by carefully examining probable impacts of the 4D tool's introduction into the firm's business processes.

Keywords: 4D management tool, design-build, workflow.

Introduction

4D visualization technology, the four-dimensional planning and scheduling that binds 3D models with their corresponding construction work schedule in its simplest form, has emerged rapidly (Hsieh et. al., 2006) in the past decades in construction and project management framework. This is mainly due to the increasing recognition from the construction industry of the benefits of using the 4D CAD applications (Kam et. al., 2003; Dawood et. al., 2003; Heesom and Mahdjoubi, 2004). The benefits of using 4D modeling mechanisms compare to the traditional tools are well described in several researches over time for its increased productivity, improved project coordination capability and optimized on-site resource utilization techniques etc. 4D construction management tools have experimentally been deployed in many projects in recent years. San Mateo Health Facility (Collier and Fischer, 1996) and Construction Director (Hsieh et al., 2006) are two brilliant examples among them.

Although many firms in construction industry realized the benefits of 4D tool, the decision-makers nevertheless are not confident enough in adopting this new tool for real project. A pragmatic application of this tool in practical business process is very much important towards its successful implementation in actual project. This research addresses this issue by intelligently developing company's existing workflow. This approach was carefully adopted towards successful application of "Construction Director", a well-defined 4D management tool developed by NTU (National Taiwan University) in collaboration with CTCI Corporation, Taiwan. Development of Construction Director has been summarized in a paper previously authored by Hsieh et al. (2006). The design of "Construction Director" takes advantages of several design patterns that could satisfactorily display and convey construction information with the 3D model and the work items on the daily schedule (Tsai et. al., 2008). In addition, the work plan in phase can be displayed and distinguished with different colors (Chang, et. al., 2007). Based on the current system (3D model by using

"Smart Plant Review" software) the tools and interfaces in "Construction Director" are developed to facilitate the four-dimensional construction simulations.

We selected a design-build project of CTCI Corporation, the largest engineering, procurement and construction (EPC) firm in Taiwan and one of the top 200 contracting companies in the world, as a case study. We experimentally implemented "Construction Director" (4D tool) in place of 3D reviewing tool in their existing workflow. SPR (Intergraph Smart Plant Review) has been used in CTCI for more than a decade. We observed impact of the use of SPR and "Construction Director" respectively on the existing workflow of the company and the proposed workflow re-engineered by us. After multiple interviews with the personnel of different engineering departments of CTCI, each step of the set-up workflow was carefully adjusted according to the management demands by minimizing changes to the existing workflow. We focused on how the firm used 3D reviewing tools in their existing workflow to facilitate the design and construction process and then proposed a customized workflow that introduced 4D tool (i.e. Construction Director). The approaches for the introduction of 4D tool into company workflow that this paper presents might very well be effective to help more professional use of this tool in real project.

Workflow Re-engineering Approach

Based on the information obtained from the interviews, we observed impact of the use of SPR and "Construction Director" respectively on the existing workflow of the company and the proposed workflow re-engineered by us. We adopted the following procedures to do that.

The interviews

We conducted interviews to relevant engineering personnel of different departments of CTCI, which were divided into two stages. At the first stage, we arranged a half day interview with the manager of the design department and obtained an overview of the existing workflow of a design-build project of CTCI. As the manager had been involved in the development of Construction Director (4D tool), we obtained his views for the new workflow and identified the redistribution of responsibilities among the departments after introducing the 4D tool in the existing workflow. Based on the information obtained from the first stage interview, we drafted two figures illustrating the existing workflow and the modified (proposed) workflow respectively. Then we conducted the second stage interview whereby two interviews were taken place.

At the second-stage interviews, the first one was to interview the manager of the project management department. We asked him to review the figures of the workflows we already had drafted and to confirm the roles and responsibilities of his department in both the workflows. The second interview of this stage was held in construction department with three personnel that included the manager, a senior engineer and the human resource engineer. We asked them to review workflow figures and to point out inappropriate part by concentrating on their relevant sector in the workflow, particularly the communication with sub-contractors, the inter-departmental interactions and other concerned issues during construction processes.

After the interviews, we incorporated their views and suggestions and carefully examined the possible impacts of the 4D tool's introduction into their business processes and then developed the final workflow.

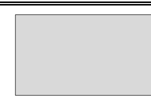
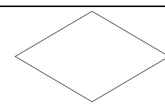
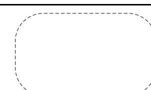
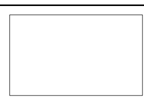
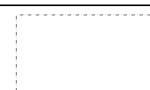
Symbols used in the workflow

Table 1 depicts the symbols used in the workflows. The arrowhead represents the flow direction of data/operation. The solid arrowhead represents one-to-many relationship and the non-solid arrowhead represents a zero relationship between the work items. We also used two dash arrowheads to represent the data exchange in either forward or backward operations. The work items are presented in rectangular boxes. The shaded rectangular boxes represent the work items associated with the 4D tool (Construction Director) that is shown only in Figure 2. The diamond box represents flow judgment. Round-edged dash line rectangle is for cooperation block that represents interaction, communication, negotiation and understanding between the departments. The sharp-edged dash line rectangle is for recursive operation that needs to be performed repeatedly until the goals have been achieved.

Existing Workflow

After the interviews, we produced a typical workflow to CTCI using the symbols explained above. As shown in Figure 1, the owner announces the initial requirements at the very outset of the project. These requirements include owner's budgetary limit, time allowance, particular objectives, expectations, limitations, and regulatory specifications etc. CTCI, as a contracting company, confirms owner's award to execute the design-build project and starts developing a project plan. Then prepares deliverables in accordance with the owner's demand and guidelines, and submit it to the owners for their concurrence. Usually owners would review the project plan proposed by CTCI and then ask CTCI to explain and redefine it several times if it does not satisfy their criteria and expectation. Upon achieving satisfaction with project plan, owners would award the final contract to CTCI to commence construction. Engineers of design department of CTCI then create a 3D model according to the plan proposed earlier to owner. Using the model, they recursively perform structural analysis and parameter study. In Figure 1, this process is presented by a recursive block. Finally, a set of necessary construction drawings are prepared to support work items and activities of the project.

Table 1 Diagrams of work items

Diagram	Representation	Diagram	Representation
	Data flow/operation (one to many)		Work item related to Construction Director(4D tool)
	Data flow/operation (zero)		Judgment
	Date exchange /operations both way		Cooperation block
	Work item		Recursive block

In the design and construction phases, CTCI usually has to lead through three major milestones in their existing workflow that includes inter-departmental meetings and owner's review for handling the design details, materials procurement and construction details. All these three landmarks need involvement of engineering personnel from different departments and concurrence of clients to effectively solve construction issues. Both the conferences and reviews rely heavily on the 3D model developed by design department of CTCI. The 3D model is usually displayed on a large screen. The concerned engineers from different departments review their individual parts regarding plan, schedule, cost and budget, site planning, manpower deployment, owner's review etc in the 3D model. Project department of CTCI usually plays coordinating role to resolve conflicts among the departments. Review of 3D model by the owner is required to achieve a final decision for further proceed. At the same time, interpreting of construction schedule and understandings of other relevant issues are needed to be clarified to the owner. Since most of the owners of CTCI do not have strong construction background, they usually hire in-house construction professionals to help them understand the project issues clearly. The owner and the project management unit monitor the quality and progress of the construction, based on the 3D model and the construction plan developed in the design phase. The site departments, including the schedule control, procurement, construction department and the sub-contractors perform construction activities according to the 3D model and the construction plan. The workflow is depicted in figure 1.

Proposed Workflow

We finally proposed a modified workflow that incorporates 4D tool (Construction Director) in the existing workflow of CTCI. The proposed workflow is presented in Figure 2. We minimized changes to the existing workflow, which will reduce the possibility of unforeseen impact on the organizational structure and on the working processes they were used to. Unlike using 3D model to review the construction plan, the engineers will use Construction Director, the 4D management tool as the major reference in the proposed workflow during their inter-departmental meetings. They also can use this tool to explain and present the plan and design results with project progress to the owners.

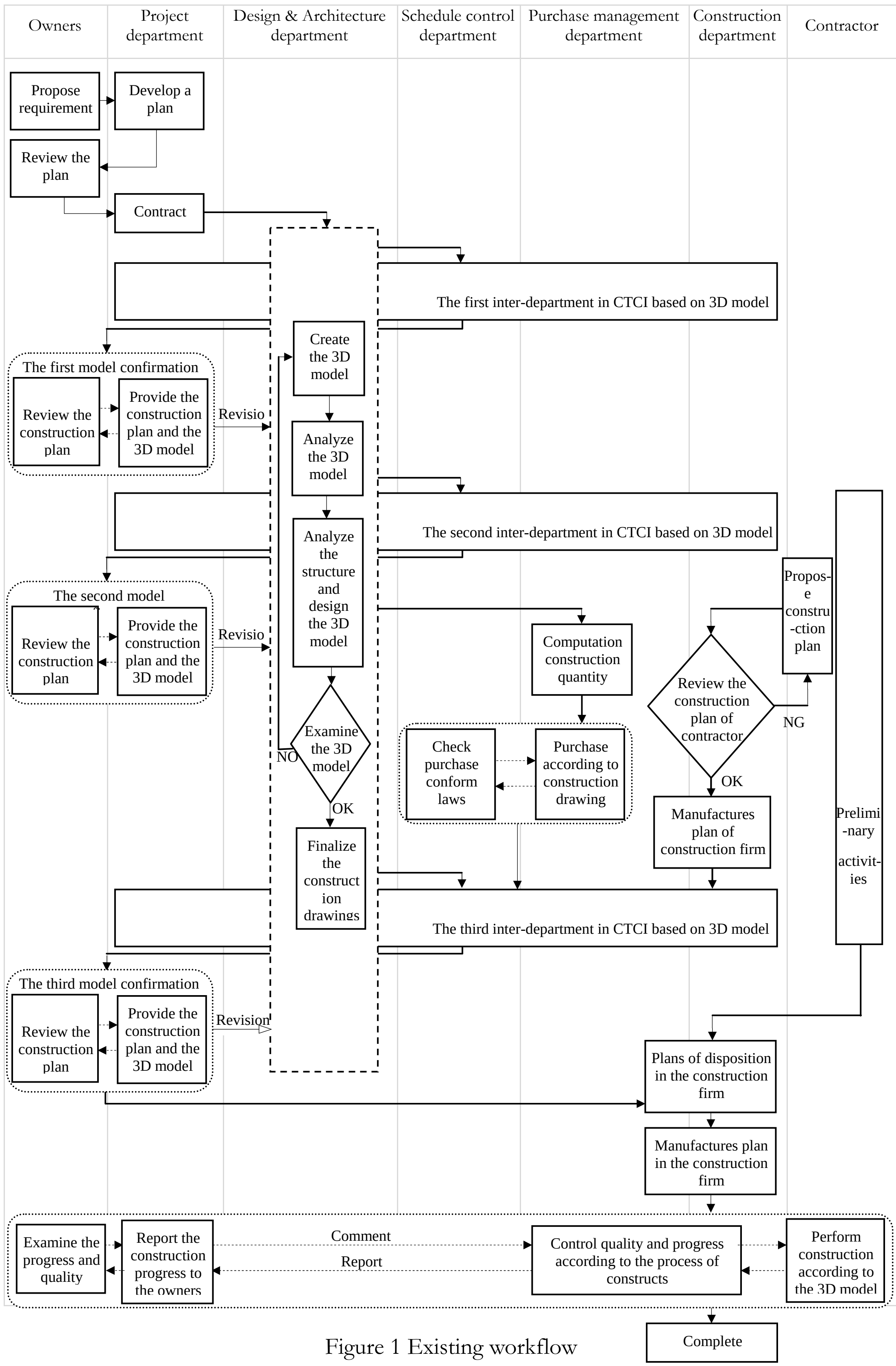


Figure 1 Existing workflow

The following paragraph summarizes the major advantages of the proposed workflow on top of the existing workflow. Firstly, the 4D model tool plays a vital role in the inter-departmental meetings of the company. These models are created by using 3D models and the schedule that presents the construction progresses visually. Therefore, 4D model can successfully reduce a large amount of time wasted for explanation and clarification of construction issues in the meetings. Secondly, 4D tool can appreciably be used in the owners' review. This will assist them in locating problems from the user's point of view and will help in making decisions with firm confidence. Thirdly, this intelligent tool can very well be used in construction sites. For a design-build project, one major concern is the communication amongst subcontractors, especially among them who work in close proximity of schedule conflicts. As well, the main contractor has to communicate and liaise with many specialized technologists, suppliers and regulating agencies with regard to execution of construction activity, material storage, access point, transport, permits and safety issues etc. It provides engineers with a more direct approach to obtain the information on hand to trace the predecessors and successors of each construction activity to avoid clashing of operations. Thus the engineering manager and the site engineers can easily take offhand decisions relating to site matter with more confidence. 4D tool also helps identify accessibility issues of movable equipment such as trucks, excavators, concrete pumps, mobile crane, gantry crane etc. and the arrangements of stationary equipment such as tower cranes, pump stations, boring machines etc. Therefore, it must improve the communication excellence between the parties involved in project works and positively enhance performances of the entire project to harmonize the operations.

Comparison between Existing Workflow and Proposed Workflow

As described in Table 2, we compare the existing workflow with the proposed workflow from two aspects. In the existing workflow engineers maintain the 3D model in their inter-departmental conferences and project review meetings. Their means of correspondence were the 3D model and the project schedule in a separate stature to devise project plan. This would be hard to facilitate engineers and the owner to understand the optimal productivity of the project prior to construction. Moreover, when different departments present their individual assignments in project coordination meetings, they ought to present the 3D model and the schedule in detail. This is a jeopardized system in communication and coordination perspective, especially if it is a large project.

In the proposed workflow, we introduced 4D construction management tool to resolve the uncertainty and ambiguity of the project matters. After introducing 4D tool, the physical constraints of site could be visualized, as well as the conflicts of concurrent activities could be identified to avoid coexistence of multi operations in one point. Therefore, the schedule controller (planning engineer) can synchronize the progress according to the accomplishment of the construction and can give immediate feedback towards organizing/reorganizing the works. And the project participants can beforehand make arrangement of procurement, equipment, manpower, exact bill of quantities etc.

According to the feedback obtained from the interviews, the operational mode of CTCI is transferred from traditional model to current 4D model. The changes showed the 3D model couldn't absolutely satisfy the company needs, thus 4D tool (Construction director) was introduced.

Conclusions

This paper presents a workflow proposed specifically for using 4D models in design-build projects. We selected a large EPC firm with an in-house 4D tool as an example case. By interviewing the key persons from different departments, we summarize the existing workflow in that firm. From the workflow, we found that the owners and engineers from different departments rely mainly on 3D models and paper-based schedule to communicate and manage the project. We proposed a new workflow for the firm, in which 4D models become facilitating tool to increase the communication between the owners and the engineers from different departments. We expect the proposed workflow can increase the degree of satisfaction from the owner, better the teamwork amongst concerned departments and eventually raise company's competitiveness in the market.

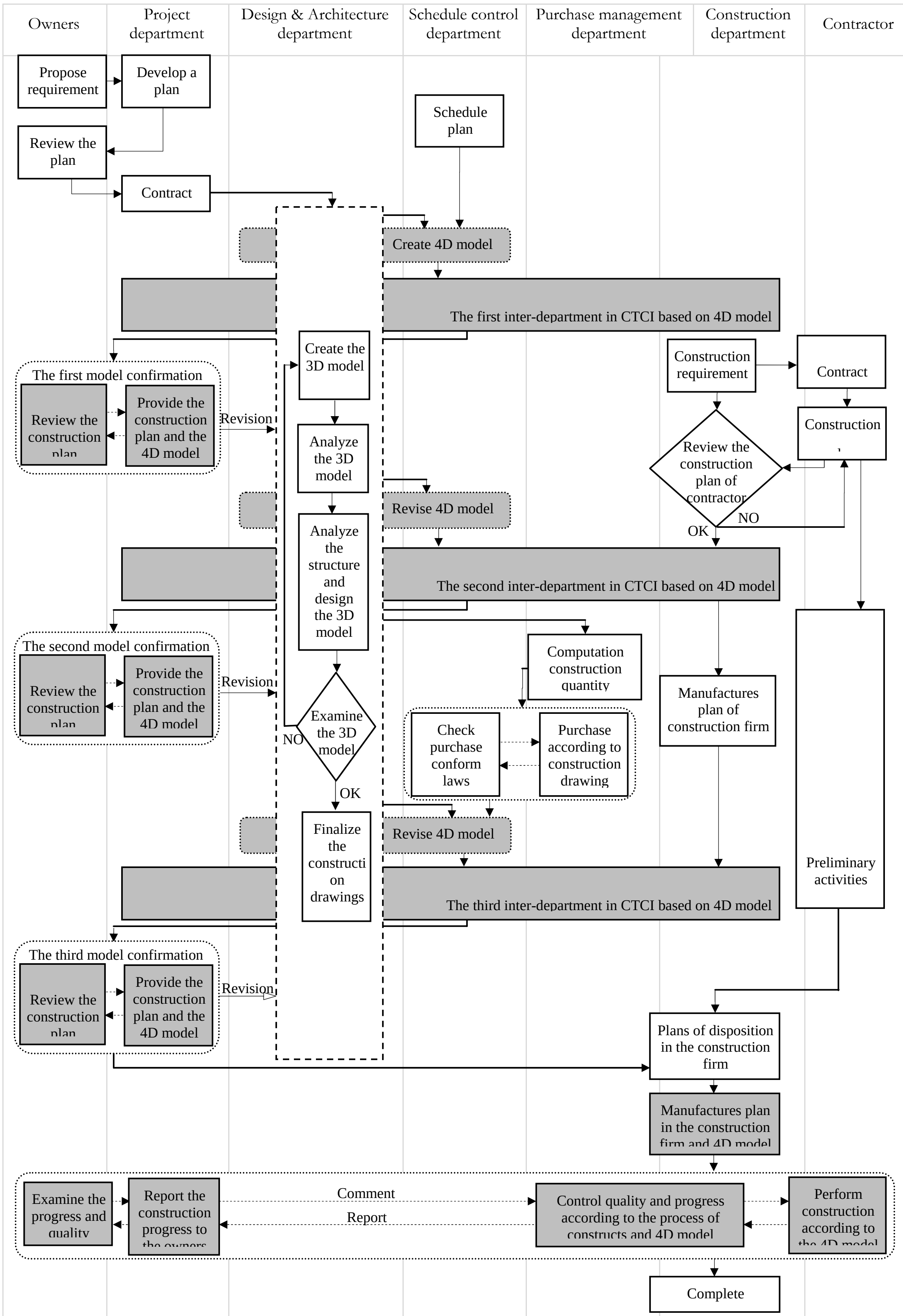


Figure 2 Proposed workflow

Table 2 The comparison of activities

Workflow	Internal Communication	External Communication
Existing workflow (3D)	1. Engineers maintain the 3D model. 2. Presenting 3D model and schedule need more activities thus require additional time to interpret. 3. Communication, coordination for team work is difficult. 4. Unable to use 3D model in different departments to know construction progress. 5. Can't efficiently plan equipment, manpower etc.	1. Owner doesn't understand project issues clearly. 2. Information gaps existed. 3. Communication gaps may exist.
Proposed workflow (4D)	1. Engineers maintain the 4D model. 2. Managing the schedule with 4D model can control complex project easily. 3. Personnel can beforehand make arrangement of equipment, manpower, procurement etc. 4. Presenting individual department's issue is easier, thus facilitates communication and coordination.	1. Report project plan with 4D model. 2. Owner can understand project related issues. 3. Visualization process helps in conveying information.

Acknowledgement

Financial support from CTCI Foundation is gratefully acknowledged. The authors would like to thank Mr. Yuan-Fu Liao of CTCI Corporation, Prof. Chuin-Shan Chen and Mr. Cheng-Han Kuo of National Taiwan University for their assistance in the development of the 4D construction management tool. We like to express our appreciation to the interviewees of CTCI.

References

[1] Chang, H.S., Kang, S.C. and Hsieh, S.H. (2007). "Color schemes in 4D construction management tools," Proceedings of Computing in Civil Engineering, Pittsburgh, U.S.

[2] Collier, E. and Fischer, M. (1996). "Visual-based scheduling: 4D modeling on the san mateo county health center," Proceedings of the Third Congress on Computing in Civil Engineering. Anaheim, CA.

[3] Dawood, N., Eknarin, S., Zaki, M. and Brian, H. (2003). "Development of an integrated resources base for 4D/VR construction process simulation and visualization," Automation in Construction, 12(2), 123-131.

[4] Heesom, D. and Mahdjoubi, L. (2004). "Trends of 4D CAD applications for construction planning," Construction Management and Economics, 22(2), 171–182.

[5] Hsieh, S.H., Chen, C.S., Liao, Y.F., Yang, C. T. and Wu, I.C. (2006). "Construction director: 4D simulation system for plant construction," Proceedings of the 10th East Asia-Pacific Conference on Structural Engineering and Construction (EASEC-10), Bangkok, Thailand.

[6] Kam, C., Fischer, M., Hänninen, R., Karjalainen, A. and Laitinen J. (2003). "The product model and fourth dimension project," Journal of Information Technology in Construction, Vol. 8, 137-166.

[7] Liao, Y. F., Hsieh, S. H., Kang, S. C., Chen, C. S., Yang, C. T., Chi, H. L., Chi, W. C. and Yu, C. H. (2007). "Development and application of 4D construction simulation tools," Proceedings of The Fourth Civil Engineering Conference in the Asian Region (CECAR), Taipei, Taiwan.

[8] McKinney, K., Kim, J., Fischer, M. and Howard, C. (1996). "Interactive 4D-CAD," proceedings of the third congress on computing in civil engineering. Proceedings of Anaheim, CA, USA.

[9] Tsai, M. H., Kang, S. C., Hsieh, S. H., Kuo, C. H., Pei, T. J. and Yeh, K. C. (2008). "Experiences on study of setting-up a 4D construction management tool in a high-rise construction project," Proceedings of 12th International Conferences on Computing in Civil and Building Engineering & 2008 International Conference on Information Technology in Construction (ICCCBE XII & INCITE 2008), Beijing, China.

Service-Oriented Integrated Information Framework for Next Generation Intelligent Construction Supply Chain Management

Tae-Hong Shin [1], Sangyoon Chin[2], Su-Won Yoon [3], Yea-Sang Kim [4], Soon-Wook Kwon[5] and Cheolho Choi [6]

[1] Dept. of Civil, Architectural, and Environmental System Engineering, Sungkyunkwan Univ, Suwon 400-746, S. Korea ; PH (82)31-290-7578; FAX (82)31-290-7570 ; e-mail:cmcic@skku.edu
[2] Dept. of Civil, Architectural, and Environmental System Engineering, Sungkyunkwan Univ, Suwon 400-746, S. Korea ; PH (82)31-290-7578; FAX (82)31-290-7570 ; e-mail:schin@skku.edu (Corresponding Author)
[3] Doalltech Co., Ltd, DMC Hi-Tech Industry Center 1580, Sangam-Dong, Mapo-Gu, Seoul, Korea 121-835 ; PH (82)2-555-9779; FAX (82)2-5555-886 ; e-mail:yoonsuwon@doalltech.com
[4] Dept. of Civil, Architectural, and Environmental System Engineering, Sungkyunkwan Univ, Suwon 400-746, S. Korea ; PH (82)31-290-7578; FAX (82)31-290-7570 ; e-mail:yeakim@skku.edu
[5] Dept. of Civil, Architectural, and Environmental System Engineering, Sungkyunkwan Univ, Suwon 400-746, S. Korea ; PH (82)31-290-7578; FAX (82)31-290-7570 ; e-mail:swkwon@skku.edu
[6] Doalltech Co., Ltd, DMC Hi-Tech Industry Center 1580, Sangam-Dong, Mapo-Gu, Seoul, Korea 121-835 ; PH (82)2-555-9779; FAX (82)2-5555-886 ; e-mail:choi@doalltech.com

Abstract

It is essential to perform an effective construction supply chain management in large-scale building construction projects. Since 2006, this research consortium has conducted a research project, named Next generation Intelligent Construction Supply chain management (NICS), that develops a process and a system that support an proactive construction supply chain management in real time by taking advantage of ubiquitous sensor network(USN) and radio frequency identification (RFID) technologies. In this project, 'next generation' means using USN/RFID technology to improve the current process of construction supply chain management, while 'intelligent' means that equipments, such as pallet, trailer, hoist, and gate, can recognize components or material that they carry or hold and communicate the information with other equipments or actors involved by using USN/RFID. This paper focuses on building an information framework that can support real time information sharing and exchange among different legacy systems operated by general contractors and suppliers as well as among equipments and actors involved. Without the information framework, information related to construction supply chain cannot be managed effectively and efficiently in the NICS environment. In the logistics industry, many researchers proposed that loosely coupled system architecture on the basis of service oriented architecture (SOA) could be a solution with providing expandability and flexibility in exchanging and sharing information among heterogeneous and geographically dispersed information systems. Based on the idea on SOA, the objective of this paper is to develop an integrated information framework and a system to support the NICS environment, where USN/RFID technologies are used, by utilizing the SOA concept. Additionally this paper describe the identification of services and information model for developing SOA based NICS environment, and include the introduction of being developed a prototype system as well.

Keywords: Information Management, SOA (Service-Oriented Architecture), Supply Chain Management, Logistics, Information Technology, Construction Management

Introduction

With the current construction industry practices aiming at building larger, super-tall structures, many new approaches are being made to ensure effective construction management, including applications of new management techniques and incorporations of IT technologies. Efficient logistics management in the construction of mega-tall buildings, in particular, is considered an essential factor that leads to a successful project. Recognizing the importance of such, wide spectrums of studies have been conducted on automated logistics management that involves the shift in paradigms in logistics management (e.g., JIT, SCM) and the use of cutting-edge IT technologies such as bar cords and RFID (Chin et al. 2005;Jaselskis 2003). Recently, logistics management in the construction industry is leaving the past's site-based delivery and release management approach for a more comprehensive system that takes care of the entire supply chain, covering orders of building materials, production, release, delivery, hoisting/heavy equipment handling, placement and installation. The expansion of the management scope is requiring all project participants, including suppliers, to adopt a new environment where all players share and co-manage necessary information on a real-time basis to guarantee successful operations. This is seen as a departure from the conventional information-processing systems that are site-or contractor-centered in nature (Chin et al. 2008). Despite the state-of-the-art technologies being adopted, the supply chain-based information management approach relies heavily on manual work, which leads to the omission or fallacy of information that contributes to the breakage in the information flow or damage to reliability due to errors in information. Other problems with the supply chain management based on the use of bar cords, RFID and other IT gadgets include difficulties in collecting and utilizing relevant information because of: (a) differences in the level of informatization by the contractor(s), the sub-contractors and the suppliers; and (b) their lack of clear understanding of the information systems or inexperience with handling such systems.

IT technologies are steadily evolving and making progress, offering varieties of techniques and possibilities. Ubiquitous sensor network (USN), in particular, is being brought closer to users, thereby suggesting various opportunities in the collection and management of information in the construction industry.

Since 2006, this research consortium has conducted a research project, named Next generation Intelligent Construction Supply chain management (NICS), that develops a process and a system that support an proactive construction supply chain management in real time by taking advantage of ubiquitous sensor network(USN) and radio frequency identification (RFID) technologies. The aim of the project is to solve problems associated with evolution of the technologies, errors made in collecting information, differences in the level of informatization shown by project-participating parties, and additional work that may become necessary to complete gathering of information. In this project, 'next generation' means using USN/RFID technology to improve the current process of construction supply chain management, while 'intelligent' means that equipments, such as pallet, trailer, hoist, and gate, can recognize components or material that they carry or hold and communicate the information with other equipments or actors involved by using USN/RFID. In other words, the NICS project is an initiative to establish an intelligent construction supply chain management (CSCM) environment where equipment, which used to be regarded as a mere vehicle to transport materials from places to places on the CSCM process, is endowed with intelligence so that they can recognize material information and can transmit and share information with other equipment via wireless networks. For instance, when a construction material such as plaster boards is manufactured at a factory and boxed, with an RFID tag attached to the box, the intelligent pallet is fed with the name of the material to allow for the release of the item. Then, the pallet reads the RFID tag and automatically recognizes the information. To release the materials whose information is loaded in the pallet, an approach is made to the intelligent trailer. Next, the Zigbee communication enables the recognition of the materials that have been loaded onto transportation for carriage and verifies whether or not correct items have been loaded for transport. Additionally, when the intelligent trailer approaches the GateSensor, the device will determine whether or not the planned materials are being released as scheduled.

As mentioned above, the present authors are developing wide varieties of supporting equipment, e.g., intelligent pallet, intelligent trailer, GateSensor (Lee et al. 2008), and intelligent hoist, as well as the systems that will effectively accommodate the automated CSCM environment. To ensure the establishment of the environment, a number of activities are being required, including the development of equipment and systems and various interfaces between equipment, between equipment and servers, or between servers and the existing legacy system. Considering in particular the industry-specific characteristics of substantial non-formality, the information management systems or solutions are called for that will be able to effectively support integration within each project or between multi-stakeholders.

Therefore, the present authors proposed a service-based integrated information framework whose features have been described earlier, and accordingly conducted processes to verify the efficacy of the framework by materializing a prototype system.

Research Methodology

The service-based integrated information framework proposed in this study refers to a comprehensive information management system wherein the information produced in stages prior to construction logistics management is efficiently collected and shared by stakeholders, as intelligent equipment are operated in a CSCM environment which adopts ubiquitous technologies such as RFID/USN. The system in question must show quick adaptability to changes in the non-formulated processes and be able to offer an integrated management environment which is based on the sharing of information via communication between various equipment items and the building of efficient interfaces with the existing legacy system.

As reviewed above, it is surely a challenge to design an information management system by effectively combining the complicated, heterogeneous and geographically dispersed environments. One solution to such task is the service-oriented architecture (SOA) concept, and a number of studies have proven that it offers an optimal approach to build an integrated management system in the logistics industry. Therefore, the present authors came up with the following processes to help develop a framework to be used in establishing a loosely-coupled integrated management environment required in NICS:

To gather and review the existing literature and case studies on SOA, based on previous NICS environment analyses;

To analyze NICS environment, by using SOA-suggested methodology(s);

To propose an SOA-based NICS management environment;

To organize services that are needed for system operations and to design a system based on the serviced identified; and

To verify the efficacy of the SOA-based integrated information framework, by materializing the prototype system.

NICS Environment

As described earlier, NICS offers a vision for the future CSCM environment and offers the establishment of automated CSCM systems centering around, as illustrated in Figure 1, RFID, USN (e.g., Zigbee, Wibro, CDMA, Wi-Fi), mobile (e.g., PDA, UMPC), and intelligent pallets and trailers, GateSensor, and intelligent hoists that show combinations of various IT equipment. At the present, equipment development for the prototype and inter-equipment communication testing are all completed.

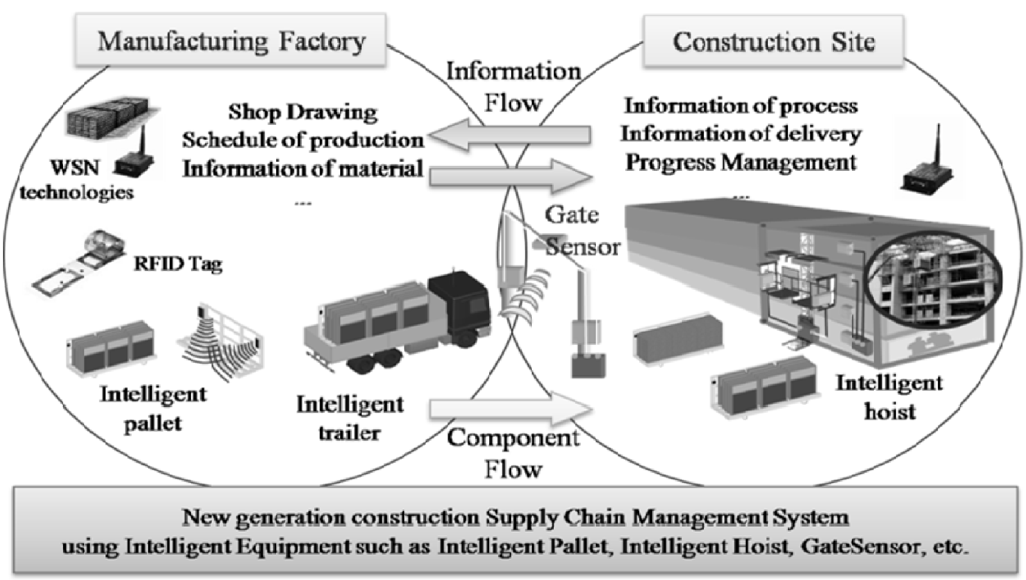


Figure 1. The vision of Construction Supply Chain Management of next generation (Lee et al. 2008)

Figure 2 below is an UML sequence diagram designed to illustrate communication between some of the intelligent equipment items, i.e., intelligent pallets and intelligent trailers.

Reviewing the diagram primarily for major information flows, the following processes are present: (1) materials are loaded onto the intelligent pallet (1.Load MaterialSet); (2) the lode cell begins to operate

(3.Recognize loaded weight); (3) this wakes up the RFID tag reader, Zigbee, and UMPC; (4) the RFID tag ID (8.Read RFID tag) recognized by the RFID tag reader is stored in the Zigbee memory (10. Write RFIDID in Zigbee memory); (5) the RFID tag ID is transmitted to UMPC's RFID tag ID via the RS232C serial communication (11. transfer RFID tag ID by RS232C); (6) the transmitted RFID tag ID works to initiate a log-in to the logistics server to pull out detailed materials information and bring it onto a display on the screen; and (7) when an intelligent pallet approaches an intelligent trailer, a Zigbee connection (17.connect Zigbee) is made, having the intelligent pallet's Zigbee memory-stored RFID tag information sent out to the intelligent trailer's Zigbee, thereby enabling an inter-equipment communication. As illustrated in the aforementioned information flow, lack of a cohesive and integrated information management system between a number of communication-supporting and display device modules and multi-servers will likely cause grave difficulties in maintaining consistency in the information. Not having a loosely-coupled interface-based system will also likely compromise the environment's quick and flexible response to changes in the modules and communication methods that comprise parts of intelligent equipment when such changes are called for. If the interface design is not a standardized one, time and cost issues will inevitably arise with each ensuing integration mission.

To help solve the problems described above, it appears necessary to build an SOA-based integrated information framework. It should be the one based on interfaces that: (a) allow for the establishment of integrated systems which efficiently accommodate heterogeneous and geographically dispersed environments; and (b) are certified as an international standard.

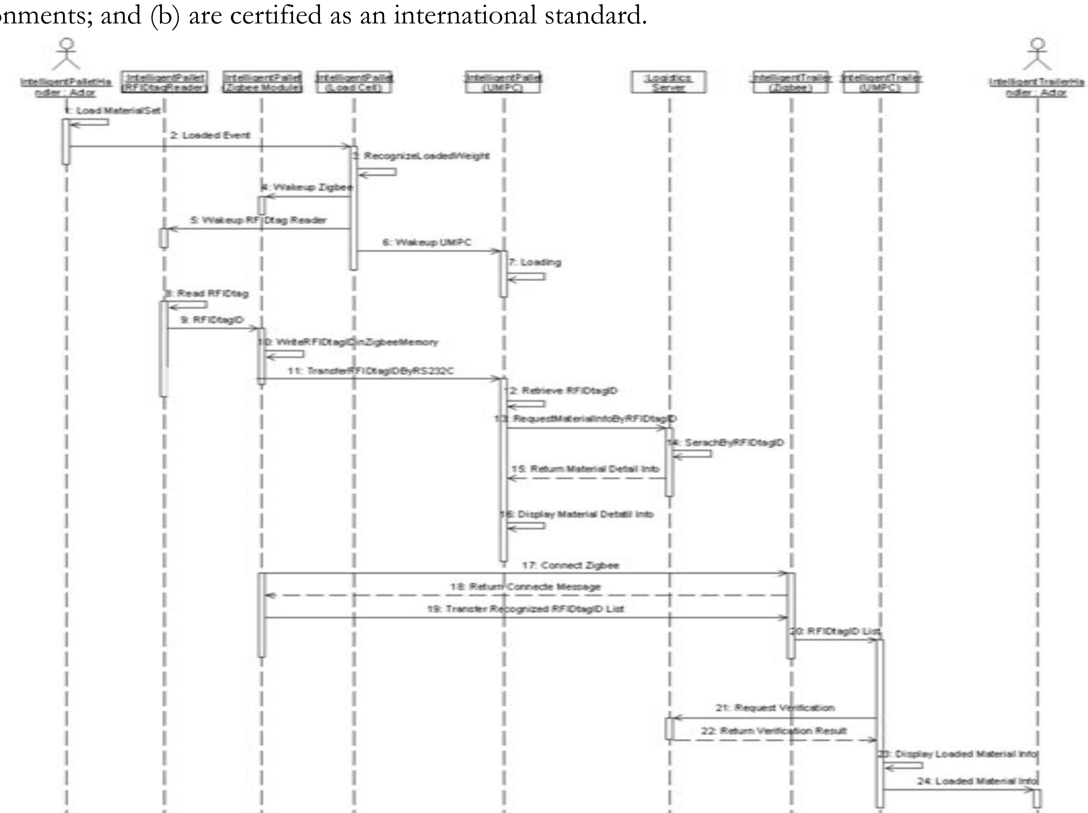


Figure 2. Communication between Intelligent Pallet and Intelligent Trailer

Review of SOA for building NICS environment

Ever since the Gartner Inc. published SOA in April 1996 as the standard interface method, SOA has been defined as a comprehensive set of IT strategies that include: (a) architecture development involving the building of web services as well as the entire application as a service unit; (b) policies to help materialize SOA; and (c) rules and shared service management. In fact, the basic concept of SOA was already being used in distribution techniques such as common object request broker architecture (CORBA) and distributed component object model (DCOM), though technical inexperience and the lack of exposure standards and cooperation between major software vendors have brought little attention to it. But with an arrival of the XML-based web services, SOA is stepping into a new limelight.

Figure 3 below shows the conceptual model of an SOA architectural style indicated in UML writing. The diagram consists of **Service Consumer**, **Service Provider**, and **Service Broker**, featuring a structure wherein **Service Description**, a standardized invoice, is used to have services that the **Service Provider** offer delivered to the **Service Consumer** by the **Service Broker** (Wikipedia 2008).

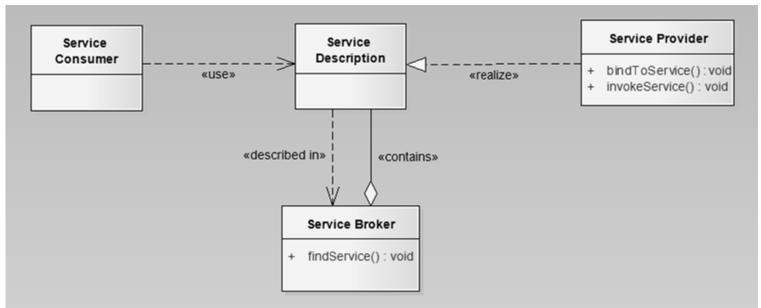


Figure 3. Conceptual model of a SOA architectural style

The biggest benefit of SOA is its ability to expose re-usable functions to the outside by complying with the standardized interface regulations/requirements and to render the functions re-usable via only service binding. The strategies also allows for the concept of software as a service, where a software program acts as a service, offering relevant functions. Recently, SOA is being applied as an architecture to help ensure integration into a legacy system, in particular to establish interface between devices or an integrated environment under heterogeneous and geographically dispersed information systems. Studies on the logistic industry have proven that effective interface between various channels on the supply chain with complicated structures is leading to fast and flexible responses to the changes in the process. Based on the findings, wide varieties of trials are being made to help coordinate and integrate complex systems during the construction of a u-City that depends on ubiquitous technologies. The NICS environment mentioned earlier is composed of multi-faceted dispersed environments that are highly susceptible to unexpected changes or variations. Given the fact, the strategic SOA approach could be effective in building an integrated management system capable of supporting the environments in an efficacious manner.

In this study, a service-based framework was established based on IBM's service oriented modeling and architecture (SOMA) methodology that offers various UML profiles at the service component level, i.e., an actual UML-based system materialization level. The approach is one of many SOA methodologies that are being offered to help design an efficient service-based integrated information framework. The Organization for the Advancement of Structured Information Standards (OASIS)'s SOA adoption blueprints were also among the alternatives. In SOMA, there is a life cycle in establishing SOA through a number of processes such as: business modeling and transformation; solution management; identification; specification; realization; implementation; and deployment monitoring and management. The present authors complied with this life cycle, though they simplified the rather complicated processes for customizing purposes, and used such new version to identify services and develop service-based applications. What is presented below is an example that indicates the fact that the development of SOA-based applications could produce desirable effects.

Let us suppose a situation where one's company is using Zigbee modules, i.e., components that are used for inter-equipment communication in an NICS environment, whose interface is designed with API base supplied by Corporation A. Due to performance and cost concerns, however, a switch to the API program offered by Corporation B appears to be inevitable. With a service-based application in place, a revision of the service that comprises the interface of Zigbee modules and a compilation process before distribution will save users cumbersome re-installation and set-up for each equipment. The entire updating work is done in only one intervention, which allows extremely fast and flexible responses to the changes. In contrast, with system materialization using non service-based formats, each equipment and program affected by the switch to a new Zigbee module API must have new installation, causing inefficiency and cost increase as well as difficulties in ensuring quick responses to a new system environment.

The proposal of SOA based NICS Environment

In order to see the aforementioned SOA concept being effectively applied to the intelligent CSCM environment, i.e., NICS, the strategies must be designed in such a way that they are able to support the CSCM life cycle between factories and sites as an integrated whole. Based on this application strategy, Figure 4 below shows an SOA-based NICS environment — an example of service infrastructure concepts that organize intelligent equipment's use and inter-equipment communication interface on the construction supply chain management process into services, as they concern interaction between factories and sites.

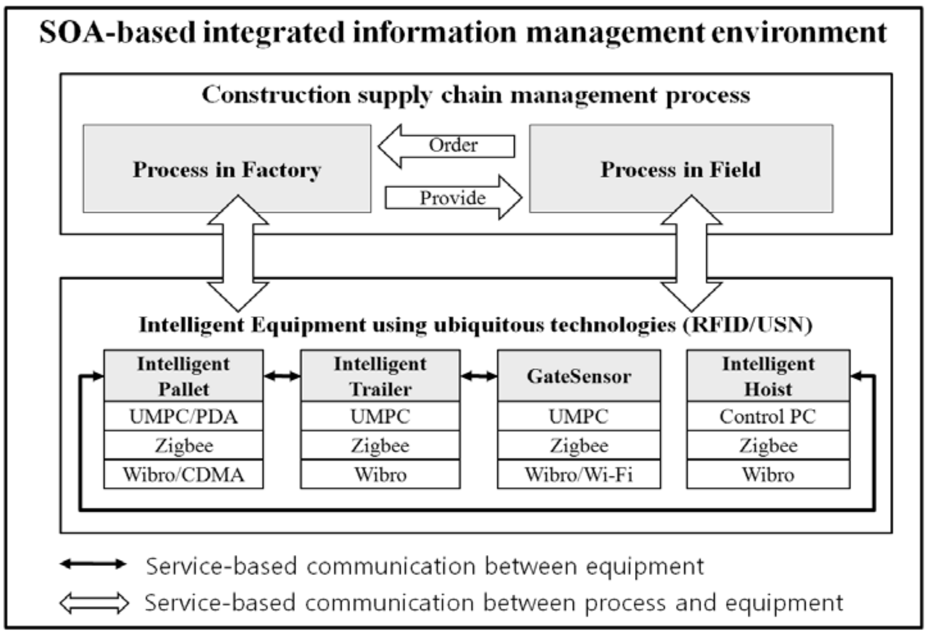


Figure 4. SOA-based NICS Environment

Service Design for building SOA base NICS environment

The service, a core component of the SOA-based environment to be established, is defined as a re-usable service in atomic form and is the smallest unit that executes loosely-coupled interface functions. Therefore, an essential question in building the SOA-based framework is how smoothly the service can be subtracted from the system.

To answer one of the questions posed in this study, i.e., to design services that are universally applicable to an intelligent CSCM environment, the present authors: (a) analyzed the process wherein intelligent equipment are mobilized and operated in the NICS environment; (b) identified the functions needed to support the process; and (c) grouped some of the functions that can be identified as being atomic and that show similar characteristics into "partitions." The services thus defined based on the aforementioned process are represented in the service component model in Figure 5's UML component diagram. As indicated in the legend at the bottom of the right-hand side of the diagram, the model consists of "components," "ports," and "expose interfaces," with "dependency" referring to the link or relationship between the components.

The total number of the organized services was 25, of which 11 partitions were formed by grouping the ones with similar characteristics. Taking a closer look at each of the partitions, main features are summarized as follows: **Code Mgmt Service Partition** manages the NICS-operated codes; **Planning Service Partition** manages the schedule software programs (e.g., primavera, msproject), PMIS/ERP, and interfaces as related to material supply planning; **Order Service Partition** supports material-ordering processes; **Delivery Service Partition** handles material transport and electronic invoicing; **Quality Service Partition** supports quality testing at factories and sites; **Progress Service Partition** helps identify the status and process of intelligent equipment use; **Notification Service Partition** supports the notification activities between stakeholders (project participants); **Communication Service Partition** supports the read/write activities of

RS232C-based serial communication, RFIDtag, and Zigbee memory stream information; **Binary Data Transfer Service Partition** supports the transmission of the images captured by the webcameras installed at the GateSensor to their servers; and **Verification Service Partition** helps identify the validity status of entering/departing trailers at the GateSensor and the material release/delivery status to verify their compliance with planned locations when intelligent hoists and pallets loaded with materials are moving to the locations. When the system is used in an actual construction environment, the interface will be materialized by referring to the interfaces whose titles begin with “I.”

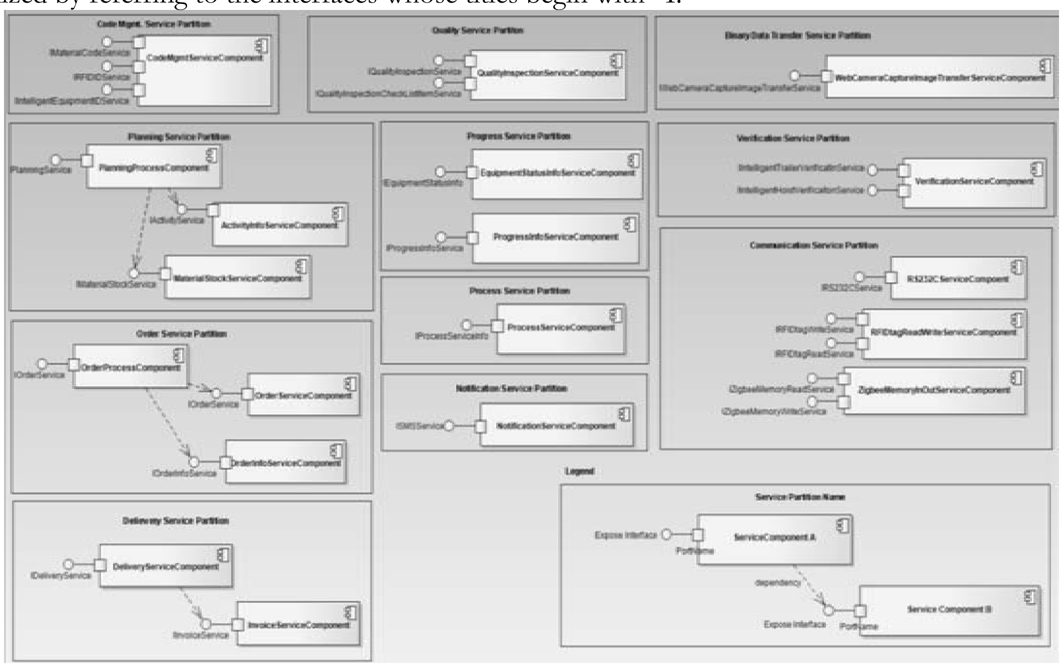


Figure 5. Service component model

SOA based NICS Prototype system

In this study, a prototype system was developed to help prove that the identified service-based integrated information framework is able to support the NICS environment effectively. For the architecture of the prototype system, OASIS’s SOA-RM (reference model), IBM’s SOA reference architecture, and Microsoft’s SOA reference architecture were reviewed to build a layered architecture, as shown in Figure 6 below, which has accommodated the characteristics of the NICS environment.

The prototype system used the windows communication foundation (WCF) which is supplied by .NET Framework 3.5 to host the core services and to effectively materialize an environment that enables inter-service transactions, with “C#” indicating codes. Figure 7 below shows a service-based application that has combined and organized the identified services as well as the intelligent trailer and pallet being used in the NICS environment, which have been built as re-usable services accommodating the functions that are needed in NICS operations via service binding and composition.

Throughout the building process of the prototype system, the present authors successfully verified that the service that is defined as an efficient interface (e.g., NICS) designed to help facilitate CSCM, which is a complicated distribution environment, is a fast and flexible model to accommodate various environmental changes (factors) as supported by a loosely-coupled comprehensive environment and that such service is effective in supporting the future CSCM environment.

Conclusions

This research aimed to suggest an SOA-based system to help establish an effective comprehensive information management environment between construction equipment and systems as part of the NICS project based on RFID/USN technologies, and to help develop services accommodating the suggested environment as well as a prototype system. To build an efficient service-based integrated information framework, IBM's SOMA method was adopted to help identify a total of 25 services which were then organized into a service component model. Next, a layer-based system architecture was designed using OASIS's SOA-RM (reference model), IBM's SOA reference architecture, and Microsoft's SOA reference architecture as bases.

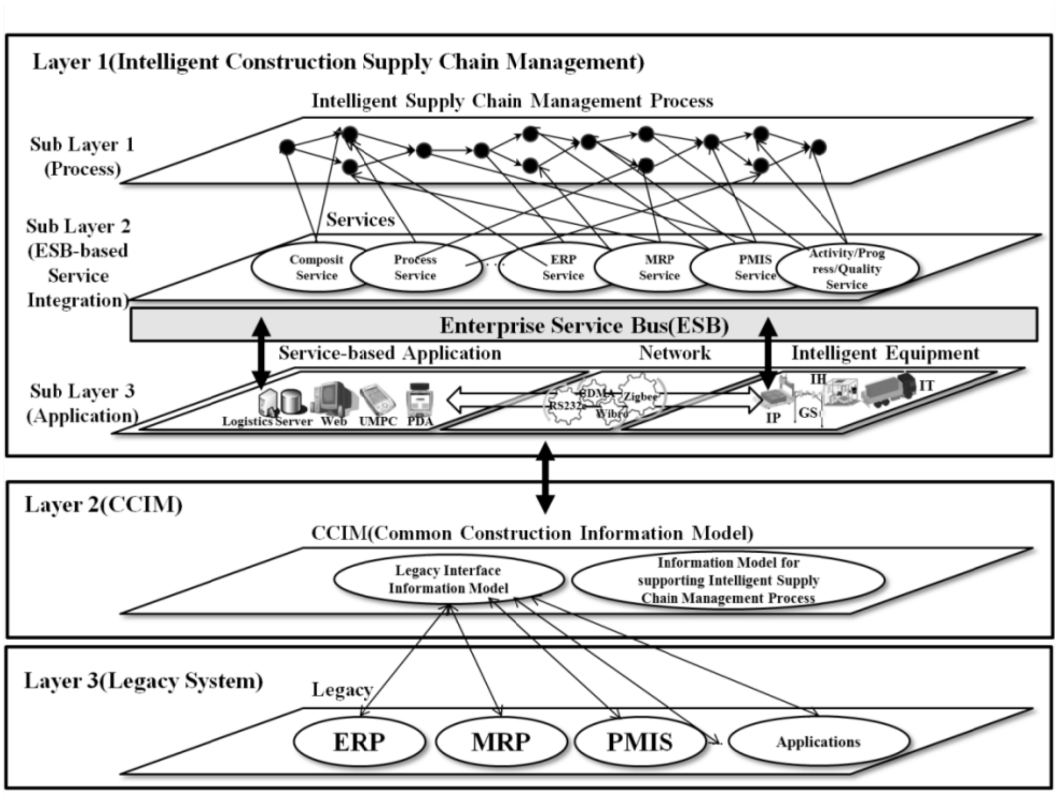


Figure 6. Layered-based System Architecture

The proposed SOA-based system environment design used the .NET Framework 3.5 WCF, which is capable of effectively supporting the service host as well as the service transaction. And a prototype system was developed to help determined whether or not the suggested design is applicable in a real-life environment, which is based on the service model and concept proposed in this study. In addition, the present authors verified that the proposed service-based information framework was a valid architecture that is able to quickly adapt to the CSCM environment where changes occur as frequently as in complicated heterogeneous and geographically dispersed information systems like NICS.

The results of this study are expected to be used in the future establishment of a full-scale NICS, as a core architecture of the system, and to help define and materialize re-usable services that come in the form of various functions required in an intelligent CSCM environment.

Acknowledgement

This work was supported by the Korean Institute of Construction & Transportation Technology Evaluation and Planning (KICTEP) with the program number of "06-Unified and Advanced Construction Technology Program-D16."

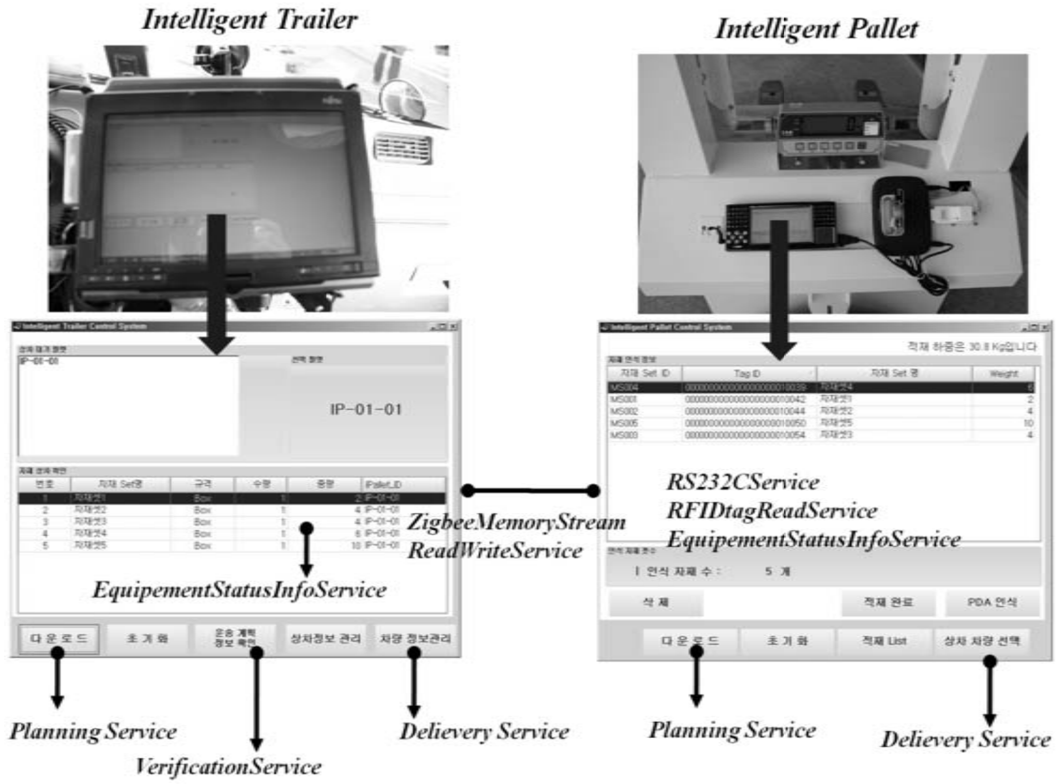


Figure 7. Service-based Application

References

[1] Li Y., and Dong, B. (2006). “Supply Chain Information Sharing System Based on SOA”, Logistics technology J., No.7, 217-220.

[2] Lv Xinyan. (2006). “Using SOA to design a public logistics service platform”, Computer systems and applications J., No.11, 33-37.

[3] Jaselskis, E. J. and El-Misalami T. (2003). “Implementing Radio Frequency Identification in the Construction Process.” Journal of Construction Engineering and Management, ASCE, New York, NY, 129(6), 680-689.

[4] Sangyoon Chin, Suwon Yoon, Yea-Sang Kim, Jeongwon Ryu, Cheolho Choi, Chang-Yon Cho .(2005). “REALTIME 4D CAD + RFID FOR PROJECT PROGRESS MANAGEMENT”, Volume 183, 33-43

[5] Sangyoon Chin, Suwon Yoon, Cheolho Choi, Changyon Cho. (2008). “RFID+4D CAD for Progress Management of Structural Steel Works in High-Rise Buildings”, J. Comp. in Civ. Engrg. Volume 22, Issue 2, 74-89.

[6] Wikipedia. (2008). “SOA”. <http://en.wikipedia.org/wiki/Service-oriented_architecture> (Sep. 27, 2008).

[7] Woo-Jae Lee, Jae-Hong Song, Soon-Wook Kwon, Sangyoon Chin, Cheolho Choi.(2008). Yea-Sang Kim, “A Gate Sensor for Construction Logistics”, The 25th International Symposium on Automation and Robotics in Construction. ISARC-2008, 100–105.

[8] Zeng HaiJin, Liu WenJu, and Ke Yong Zhen. (2008). “Supply Chain Simulation : Collaborative Design System based on SOA – A Case Study in Logistics Industry-“, Proceedings Of World Acadamy Of Science, Engineering and Technology Volumn 29.

Automated Building Information Modeling System for Building Interior to Improve Productivity of BIM-based Quantity Take-Off

Seong-Ah Kim [1], Sangyoon Chin [2], Su-Won Yoon [3], Tae-Hong Shin [4], Yea-Sang Kim [5] and Cheolho Choi[6]

[1]Dept. of Civil, Architectural, and Environmental System Engineering, Sungkyunkwan Univ, Suwon 400-746, S. Korea; PH (82)31- 290-7578; FAX (82)31- 290-7570 ; e-mail : Kody25@skku.edu
[2]Dept. of Civil, Architectural, and Environmental System Engineering, Sungkyunkwan Univ, Suwon 400-746, S. Korea ; PH (82)31- 290-7578; FAX (82)31-290-7570 ; e-mail : schin@skku.edu (Corresponding Author)
[3]Doalltech Co., Ltd, DMC Hi-Tech Industry Center 1580, Sangam-Dong, Mapo-Gu, Seoul, Korea 121-835 ; PH (82)2- 555-9779; FAX (82) 2-5555-886 ; e-mail : yoonsuwon@doalltech.com
[4]Dept. of Civil, Architectural, and Environmental System Engineering, Sungkyunkwan Univ, Suwon 400-746, S. Korea ; PH (82) 31-290-7578; FAX (82) 31-290-7570 ; e-mail : cmcic@skku.edu
[5]Dept. of Civil, Architectural, and Environmental System Engineering, Sungkyunkwan Univ, Suwon 400-746, S. Korea ; PH (82) 31-290-7578; FAX (82) 31-290-7570 ; e-mail : yeakim@skku.edu
[6]Doalltech Co., Ltd, DMC Hi-Tech Industry Center 1580, Sangam-Dong, Mapo-Gu, Seoul, Korea 121-835 ; PH (82) 2-555-9779; FAX (82) 2-5555-886 ; e-mail : choi@doalltech.com

Abstract

Recently Building Information Modeling (BIM) has been widely used to manage building information throughout the project life-cycle more effectively and efficiently. Particularly in quantity take-off and estimation, BIM-based process is getting more and more attention, and even BIM-based quantity take-off at the most detailed level has been performed in several building projects in South Korea. The practitioners involved in those projects have pointed out that modeling rough and finish interior of a building occupies a big portion of time in the whole modeling process and the manual modeling for interior is error-prone, which could cause serious result, such as wrong cost estimation and disputes. To resolve the problem, this research proposes an automated modeling method that model a building interior automatically after selecting an interior method by a room or space basis. This research develops a interior modeling method and a system that can model various rough and finish interior components automatically at once after a user select the type of interior for a given space or room. To do so, this research identifies typical interior types of buildings and a house built in South Korea, and develops a mechanism that can support modeling both typical and non-typical types of interior by allowing the flexibility in selecting interior material and components and the order of construction process for the selected items. The automated modeling system for building interior has been preliminarily tested at a typical condominium building project and found that the system could improve dramatically the productivity of BIM-based quantity take-off and estimation process.

Keywords: Building Information Modeling (BIM), Three-dimensional CAD model, Quantity Take-off

Introduction

As Building Information Modeling (BIM) has spread in recent years, many three dimensional Computer Aided Design (CAD) commercial programs used in varied fields such as design, structure, equipment, execution of construction work and estimation have come onto the market. In particular, an application that enables automated quantity take-off from 3D models in the estimation part is examined. BIM based quantity take-off and estimation have been executed by using the application in many projects in South Korea. Also, it is recognized that 3D based estimation applications are better than 2D based ones in terms of accuracy of quantity and practical use of 3D models. However, it has been discovered by practitioners who worked in the projects that BIM based estimation process needed to be improved while the projects are in progress. One of the factors that leads to that improvement is the task time of the process of modeling the materials used in the late stage of building increases when rough/finish interior modeling for the purpose of estimating the precise quantity of materials.

Nowadays, sketches in South Korea construction projects use 2D drawings, therefore, it is necessary to produce 3D models from2D drawings for 3D based quantity take-off. Also, it has been observed that total working time of 3D based estimations take twice longer when using an old estimation system, as 3D modeling takes as much time as estimation does. In addition, it reveals that an effective modeling method is critical since interior finishing modeling takes 90 percent of the total time due to many kinds of materials used. In the case of South Korea, estimation needs a rapid reaction as well as the level of requirements in terms of 3D modeling; otherwise, it becomes a deterrent to induce the country to use 3D modeling.

Hence, this research will recommend a scheme that improves productivity via an automated method that allows solving the problem of increasing total working time when the process of modeling many sorts of materials.

Characteristics of 3D model for quantity take-off

3D based quantity take-off represents the extraction of quantity data such as length, size, and volume related to quantity from 3D CAD models and connecting the data to the content. Furthermore, this meaning is based on the underlying assumption that materials in the content should be produced for the 3D model or object because there will be no possibility of extracting quantity information without a 3D model.

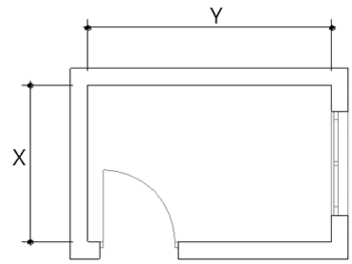
In the stage of detailed estimation to get the closest quantity to the quantity in real construction projects, exact small quantities for each material are calculated. In order to compute 3D based quantity take-off in the stage of detailed estimation, the modeling process that produces 3D models for each material constructed is compelling. With the exception of 3D modeling for small supplies such as nails and screws, other major materials have to be modeled according to the content so that the quantity of the content connected with the related materials can be measured.

3D model in the design stage consists of essential elements that determine the room and building exteriors. However, it is very common that this stage does not involve dealing with materials for interiors. Furthermore, using the accomplished model from design stage is impracticable so as to measure the quantity take-off because there are some differences between the 3D model for the quantity take-off and building designs. It is possible to use 3D models from the design stage in the rough estimation stage without any changes. However, it is unworkable to apply it in the phase of calculating detailed small quantities for each material and they needed to be modeled so as to measure the precise quantity take-off.

Limitation of 3D based quantity take-off for construction finishing

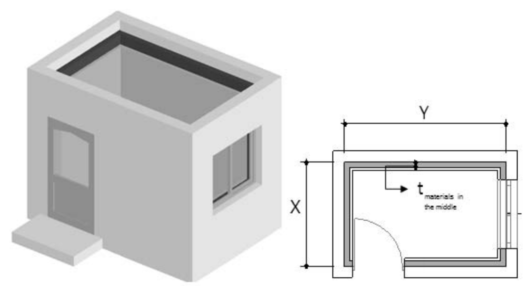
2D based estimation figures out the quantity for each material after referring to the room, measurements coming within the purview of width and length and types of materials for interiors, and then those can be combined detailing the amount of construction expenses. The size coming under the room width and length is standardized according to the interior size of the bony frame, X and Y from figure 1. Also, it is viable to apply the same width to the content and compute the quantity without the classification between materials for the middle phase and the final stage. As materials for the middle phase and the final stage are modeled in the same way as the real construction state from figure 2, however, the quantity of materials in the very last phase is different depending on the thickness of the materials in the middle stage. Moreover, the finished construction model for 3D based quantity take-off enables not only to estimate the precise amount but also to examine viability of the construction.

For the accurate quantity take-off, materials have to be modeled in the exact state of construction, and, without modeling, it is possible to refine the 3D model in order to modify wrong quantity measurement. While modeling the materials as well as the bony frame for the purpose of 3D based quantity take-off for construction finishing, the increase of the task of reforming the model is caused by the increase of the range of modeling. It is feasible that 3D based quantity take-off is admitted in terms of quantity accuracy, on the other hand, it is not efficient and widely applicable in terms of the quickness of the task because the process needs to be formed in 3D model which causes the delay of the task.



Products	Types of materials	Factor of mesurement
Cementmortar W18	materials in the middle	$2\times(X+Y)\times h_{\text{materials in the middle}}$
Water Paint	materials in the final	$2\times(X+Y)\times h_{\text{materials in the final}}$

Figure 1. 2D based estimation example



Products	Types of materials	Factor of mesurement
Cementmortar W18	materials in the middle	$2\times(X+Y)\times h$
Water Paint	materials in the final	$2\times(X+Y-2t_{\text{materials in the middle}})\times h_{\text{materials in the final}}$

Figure 2. 3D based quantity take-off example

Automated Building Information Modeling System for Building Interiors

Generally, creating 3D construction modeling consists of elements that determine the shape of rooms. They include historic walls, supports, sleeves and beams which are represented by the building structure and also non-historic walls, blocks and simple frame walls that are used for partitioning. 3D models enable showing similar images of the real construction since it is possible that the model present the quality of the materials. Consequently, it is not viable to take a look at the detailed shape of the materials cutting a section of the 3D model completed at the real stage. This partially results from the fact that it is not common to model every single material at a time in the modeling part, furthermore, it is likely to happen that the single section of the wall including the materials is modeled by an object. It grants to reduce the time of the task to be efficient.

To compute 3D based quantity take-off for construction finishing, it is required to create the 3D model by each type of material with the aim of measuring the quantity as precise as the real state of construction in the detailed part. Also, it is needed to take an action to develop the productivity of the estimation tasks due to the increase of modeling time and the amount of the tasks for the 3D based quantity take-off. Hence, this research formulates the automated building information modeling system to trim down the progress of modeling the materials required and reduce the time of building modeling tasks.

It is very reliable to state that automated building information system enables practitioners to cut the iterated tasks of 3D modeling for the quantity take-off and produce automatically the materials according to the shape of the frame. In other words, it allows practitioners to reduce the amount of the task and the time by using automated system which make internal modeling possible instead of the old method that practitioners repeatedly draw each material followed by the baseline of the frame.(figure 3.)

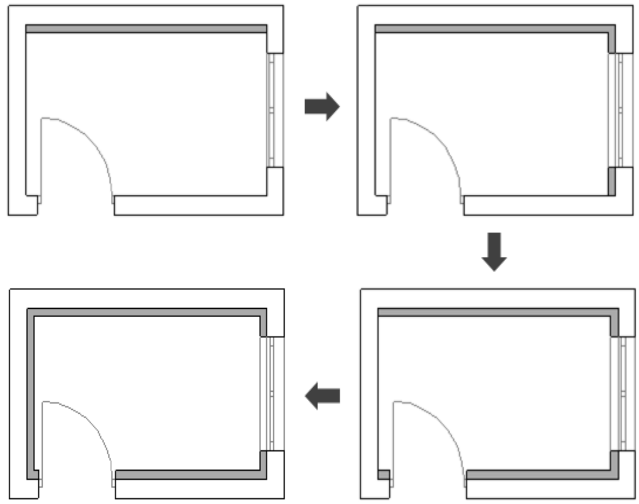


Figure 3. the old method of creating 3D model

Automated Building Information Modeling System Developing and Inspecting

As indicated, in the old way, practitioners used to draw many materials over and over to create 3D models. Hence, this study has developed an automated building information modeling system to execute drawing tasks internally rather than practitioners drawing manually so that the productivity of the tasks is improved and total time of 3D modeling is reduced.

In the way that practitioners draw manually, the sketches are also decided by their own thoughts; however, it is necessary to interpret the frame area and recognize the baseline for the aim of modeling the materials in the system. Accordingly, the frame of the building up to standard should be set as geometric constraints when modeling materials. The system internally seeks the connection in the frame that users selected and calculate room or coordinates in order to identify the baseline of the finished modeling. Furthermore, it checks the possible spots of the opening and applies to the final stage of the modeling. (Figure 4.)

This system has performed a pilot test which makes it an object of the typical interior modeling of common houses widely used in South Korea so as to examine the result of the system from this study. By testing the system, it has been outlined that the total number of promising modeling is 84 of 211 which is approximately 40%, and about 50% of the 40% which is 51 of the materials needed to be improved. Also, 33 materials from the result of the system could be used without any modifying and the time of modeling has been reduced to about 60% to 70% compared to manual modeling.

Conclusions

The research includes the criteria as needed in the stage of detailed estimation for 3D model based quantity take-off and its limitations. To overcome the limits, the finished modeling method that allows to draw many materials at once on the basis of geometric constraint and the Automated Building Information Modeling System based on it have been developed. The system has been trialed to the typical interior modeling of common house which is widely used in South Korea, and, in accordance with the test, higher productivity in terms of 3D based quantity take-off and estimation has been discovered due to the time reduced.

References

[1] K. Nassar, W. Thabet, Y. Beliveau. (2003). "Building assembly detailing using constraint-based modeling", Automation in Construction, Volume 12, Issue 4, 365-379.

[2] Rafael Sacks, Charles M. Eastman, Ghang Lee. (2004). "Parametric 3D modeling in building construction with examples from precast concrete", Automation in Construction, Volume 13, Issue 3, 291-312.

[3] Hilderick A. van der Meiden, Willem F. Bronsvoort. (2007). "Solving topological constraints for declarative families of objects", Computer-Aided Design, Volume 39, Issue 8, 652-662.
[4] Chia-Hui Shih, Bill Anderson. (1997). "A design/constraint model to capture design intent", ACM Symposium on Solid and modeling and applications, 255-264.
[5] Javier Monedero. (2000). "Parametric design: a review and some experiences", Automation in Construction, Volume 9, Issue 4, 369-377.

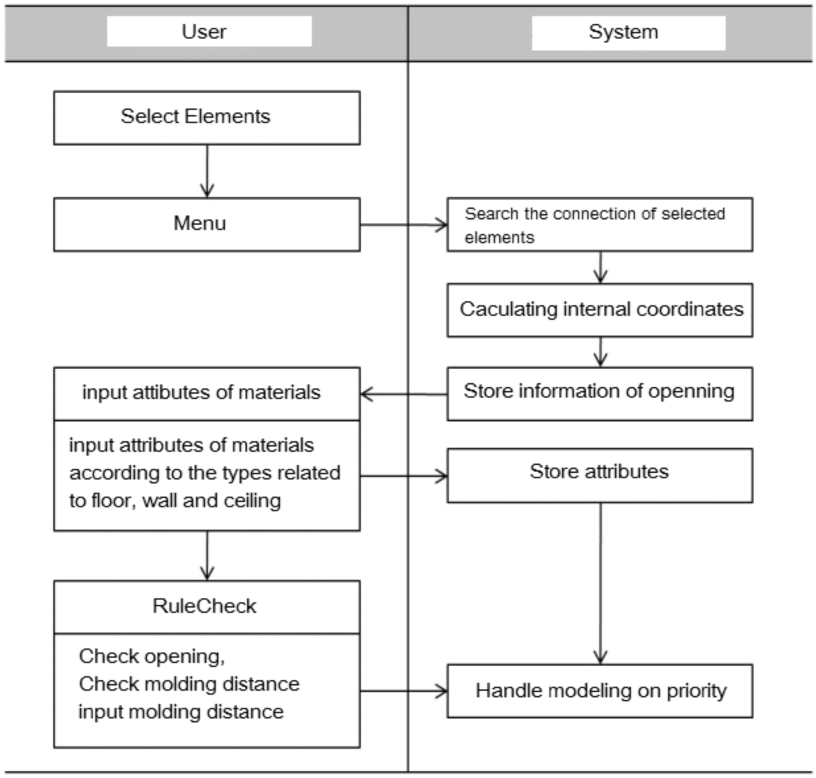


Figure 4. Automated Building Information Modeling System Process for Building Interior

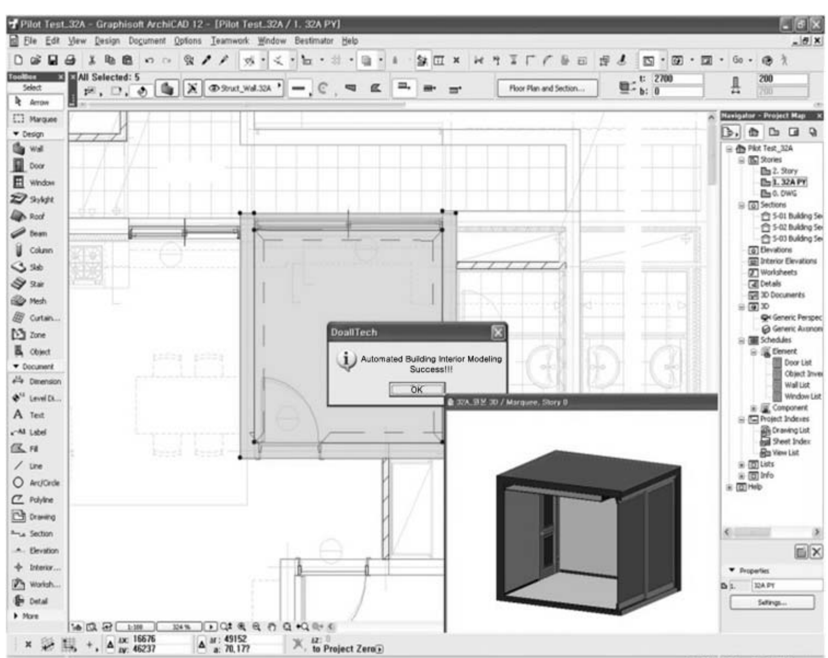


Figure 5. Pilot Test for Automated Building Information Modeling System for Building Interior

Automation of ROAD Maintenance – Development of a Roughness Measurement System for the Quality Control of Gravel Roads

Jussi Seppä [1] , Rauno Heikkilä [2]

[1]Construction Technology Research Center, Structural Design and Construction Technology Laboratory, Department of Mechanical Engineering, University of Oulu; P.O.Box 4400 FIN-90014 Oulu, Finland; PH +358 403586015; FAX +358 8 5534322; e-mail: jussi.seppa@oulu.fi
[2]Construction Technology Research Center, Structural Design and Construction Technology Laboratory, Department of Mechanical Engineering, University of Oulu; P.O.Box 4400 FIN-90014 Oulu, Finland; PH +358 40 5385840; FAX +358 8 5534322; e-mail: rauno.heikkila@oulu.fi

Abstract

The main goal of this research was to develop an automatic roughness measurement system for quality control of gravel roads. The measurement system is based on the vertical acceleration sensor which is placed on the rear right suspension arm of the vehicle. The measured accelerations are collected to a terminal box which processes the data and sends it wirelessly via Bluetooth to information management software. The software runs in a mobile phone and operates with GPS. Classification of road condition is shown in real time on mobile phone screen. The measurement details are sent to server where all measuring information is available to specified user group.

Keywords: acceleration sensor, gravel road, roughness measurement, wireless data transfer

1. Introduction

Up to today the most common method to estimate gravel road roughness is visual observation or IRI (International Roughness Index) measurements. However, visual observation has been found to be insufficient for exact and accurate quality determination and the results are always dependent on individual surveyors. In addition, measured and computed quantities for IRI values do not give enough information what kind of roughness we in practice have. Previously research has shown that measured vertical acceleration can be better way to measure roughness of gravel roads than IRI measurement. (Lampinen 1998, 2001)

The roughness on gravel road affects driving costs and riding comfort and can be even unhealthy. The roughness increase vehicle costs such as tire and fuel cost. The roughness of road creates mechanical vibration which affects human concentration and motivation. The vibrations which are near the resonance frequency of a human's organs and body parts have the most serious influence to human. (ISO 2631–1, Marjanen 2002)

Figure 1. An example of gravel road in Finland.

The main goal of the project was to develop impartial, efficient and road-safe measurement system which could be used in the quality control of gravel roads. More specifically, the aim was to research how this measurement device could be used in different types of cars and how tire pressure affects to measurement data. The main parts of measurement system are an acceleration sensor, a terminal box and information management software. The acceleration sensor was placed on the rear right suspension arm of the vehicle. The terminal box processes the measured data and sends the data via Bluetooth to the information management program which runs in mobile phone.

2. Development of a roughness measurement system

The roughness measurement system development work consisted of two main parts. The first part was research of an acceleration sensor measurement which is based on vertical accelerations of the rear right suspension arm of the vehicle. The second part was to develop the data transfer from the measurement unit to the information management program called T&M Autori. T&M Autori software runs in Nokia mobile phone and operates with GPS.

2.1 Measurement equipments

Roughness measurement system is based on the capacitive acceleration sensor (measure range 12 g) which is placed the rear right suspension arm of the vehicle (figure 2). The sensor location minimizes the effect of suspension on the measurement result. The sensor measures vertical accelerations which are caused by roughness of road. The sensor must be installed in the right position, so, there there won`t be any extra measurement error. When the sensor is installed in the right position it measures one g acceleration. This one g acceleration changes a bit when driving the uphill but the effects on measurement result are not significant. Roughness of gravel road typically causes low-frequency vibrations to vehicle structures. It is efficient and economical to measure these oscillations by an acceleration sensor.

Figure 2. The acceleration sensor is placed on the suspension arm of the vehicle.

The information management program and acceleration sensor were developed separately, i.e. in the development stage of roughness meter there wasn't joint operation between the information management program and the acceleration sensor. The acceleration signal was logged to a PC by an oscilloscope. That way it was possible to research measurement more closely.

The measurement equipment also includes a terminal box (figure 3). In the development stage of the roughness meter the terminal box was not required because measurement information was logged to PC. In practice, the processor of the terminal box will later process the acceleration signal and will send information via Bluetooth to the mobile phone software.

Driving speed seriously affects seriously to measured acceleration, and must be taken into account. All measurements are modified to a reference speed v_{norm} which is 60 km/h. This is the most common speed on gravel roads and typical measurement speed is usually close to 60 km/h. It is important to note that the sampling rate is constant. This results in the following formula.

$$k = \frac{v_{norm}}{v_m}, \qquad (1)$$

where k = change factor, v_m = vehicle measurement speed.

The mean value (absolute) and RMS (Root Mean Square)-value are calculated from acceleration signal by the equations 2 and 3. Values are calculated 1.2 seconds apart, because these 1.2 seconds parts are same size than the information what terminal box will send to mobile phone software. In the development stage of roughness meter these values were calculated in post-processing and measurements were made only at the standard speed.

$$|\bar{x}| = \frac{1}{n}\sum_{i=1}^{n}|x_i| \qquad (2)$$

where x_i = measured observation, n = number of observations

$$RMS = \sqrt{\sum_{i=0}^{n}\frac{x_i^2}{n}}, \qquad (3)$$

Figure 3. The terminal box.

2.2 Measurements

During development there were several field tests during year 2008. Different kinds of unevenness were measured to found out what kind of accelerations they cause. That way we can also research about limit values of gravel roads condition rating. Measurements were made with two different vehicles (Citroën Xsara Picasso and Volvo V70 2.5T AWD) and with different driving lines.

Another test was made with a damper testing device and the above-mentioned acceleration measurement device. The vehicle damper testing device is used in vehicle inspection at least in Finland. The damper testing device operating principle is simple. The vehicle front or rear tires are driven on a test plate which gives each tire a couple of seconds' vibration. In this case the main goal was to research how a tire pressure affects measurement results and is there any difference between two tire types. Another goal was to research how this damper testing device is applicable to calibration of the acceleration-based measurement method.

All measurements are done with a 1 ms sampling rate on a PC-oskiloscope program. The driving speed was kept as uniform as possible. The speed varied 40-60 km/h but all result are transformed to reference speed 60 km/h. using factor k (equation 1).

2.3 Development of wireless data transfer

The terminal box sends measurement information every 1.2 second to information management program (T&M Autori) via Bluetooth. The main task in development work was to develop an interface between the terminal box and mobile phone software. The terminal box will send processed measurment information to software. T&M Autori software works in Nokia mobile phone and operates with GPS (figure 4). The software correlates speed and measurement data; therefore vehicle measurement speed can be variable. T&M Autori computes roughness parameters for every 100 meters. The gravel road evenness information could be sent to a server for later review and reporting. (Tietomekka 2008)

Figure 4. Wireless information gathering and recording system. (Nokia Communicator, GPS, Bluetooth, 3G).

The data management software will produce a limit value for every 100 meter. There will also be a calibration option so that it will be possible use acceleration measurement device in different vehicles. The Autori software has been tested with the acceleration device. The final report of the measurement results has been developed. The Autori will send the information via wireless to a server where the final measurement result can be visualized.

3. Results

3.1 Measurements

Figure 5 shows the acceleration signal from Finnish a gravel road. The horizontal axis is time and the vertical axis is acceleration. The measured road part was 200 meters long. The first 100 meter part was in good condition, but the end of the road was much more uneven. There were also plenty of potholes. The acceleration signal level is quite low in the first part of road. Between 7.5-9 seconds there are huge accelerations which are caused by potholes.

The same road part measurement result are shown on table 1. The time means time from the start of measurement. The acceleration average is much higher on the second part of measured road. The road roughness clearly affects the measurement result, as desired.

The results of measurements using two different vehicle types are shown on table 2. There is a noticeable difference between these two vehicles results. The results which are measured on the Citroën are higher than those on the Volvo. In this research the different driving line affects also to the measurement results. The results are bigger from the right border of the road than from the middle of the road centre line.

The results of tests wich are made with damper testing and acceleration measurement devices are shown in figures 6 and 7. There are no clear differences between the two tire types (normal, or with spikes). There seems to be a correlation between tire pressure and acceleration value. As a rule, when tire pressure grows the result of acceleration measurement also grows also a little bit.

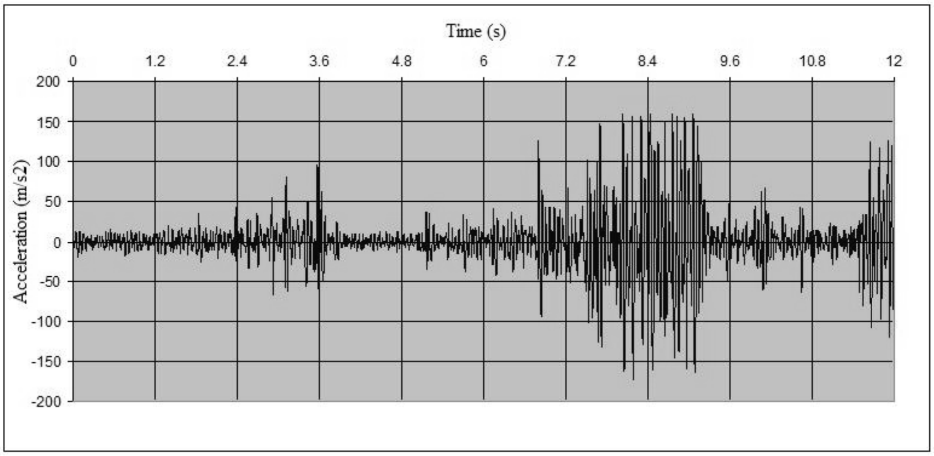
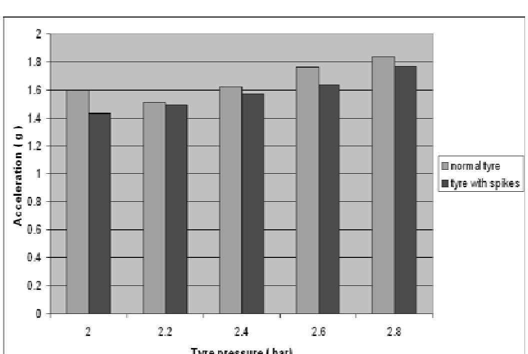
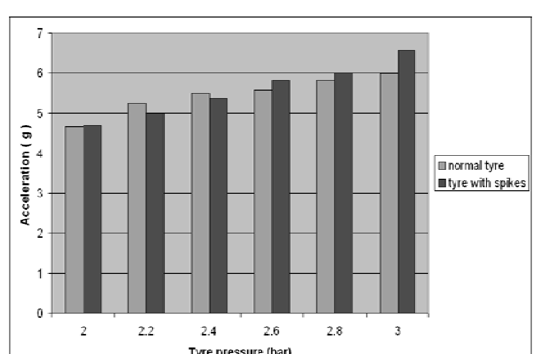


Figure 5. Acceleration signal from finnish gravel road.

Table 1. The measurement result from Finnish gravel road.

Time [s]	Acceleration average (absolute value) $[m/s^2]$	RMS $[m/s^2]$	Acceleration average of 100 meter $[m/s^2]$	RMS average of 100 meter $[m/s^2]$
0–1,20	5.02	6.23	**8.95**	**12.43**
1,21–2,40	7.74	9.73		
2,41–3,60	16.67	23.59		
3,61–4,80	6.67	11.16		
4,81–6,00	8.57	11.45		
6,01–7,20	17.26	24.53	**30.26**	**41.78**
7,21–8,40	44.1	59.02		
8,41–9,60	52.55	69.18		
9,61–10,80	13.64	19.03		
10,81–12,00	23.72	37.18		





Figures 6 and 7. Left: Tire pressure effect to 1 second acceleration average value. Right: How normal and winter tire air pressure affect to maximum acceleration value.

3.2 Automatic roughness measurement system for quality control of gravel roads

The developed automatic roughness measurement system consists of parts which are shown in figure 8. The measurement information is sent to Autori-program for every 1,2 second from the terminal box. The software computes absolute average and RMS value for every 100 meter road part and compares the values to limit values. The limit values are set on different roughness of road. The result of quality control is shown in real time on a mobile phone screen (figure 8). The measurement details are sent to a server where all measurement information is available to a specified user group (figure 10). The measurement results are shown on a map based window.

Table 2. Two different vehicles measurement result from same road part. The measurement was made on three different driving lines with both vehicles. Driving line: 1 is right border of road, 2 is middle of right traveled way and 3 is middle of centre lain.

Vehicle	Driving line	Measurement speed [km/h]	Acceleration average to 100 meter $[m/s^2]$	Acceleration average of 150 meter (absolute value) $[m/s^2]$	RMS average of 100 meter $[m/s^2]$	RMS average of 150 meter $[m/s^2]$
Citroën	1	40	18,90	17,96	25,63	23,53
Volvo	1	40	13,10	12,13	17,28	15,91
Citroën	2	40	16,44	14,27	21,25	18,33
Volvo	2	40	12,30	10,17	15,48	12,91
Citroën	3	40	14,41	12,83	17,99	16,14
Volvo	3	40	11,31	10,03	14,20	12,61
Citroën	1	50	17,88	16,98	24,37	22,97
Volvo	1	50	12,80	12,29	16,40	15,70
Citroën	2	50	17,14	15,00	21,66	19,03
Volvo	2	50	11,66	10,45	14,64	13,19
Citroën	3	50	15,46	13,81	19,41	17,34
Volvo	3	50	12,42	11,33	15,72	14,51

There are some rules and guidelines to this measurement method. When the measurement device has been installed on the vehicle, calibration should be done. The calibration can be done with the vehicle damper testing device. The measurement minimum driving speed is set to 30 km/h. If the measuement speed is slower than 30 km/h there will be an error message in the measurement information. It is also recommended that the tire pressure will always be near the design value.

4. CONCLUSION

Experiences of the developed measurement method are promising. This kind of automatic roughness measuring device makes gravel roads' quality control easier and more reliable. The objective measurement is automatic and allows agencies to set quality levels for gravel roads maintenance.

The developed measurement method allows to efficient information collection automatically from gravel roads. The measurement operator no longer needs to be a professional. Now it is possible to store large quantities of evenness information and make different analyses. In future there will be also historical information available, which allow totally types of new analysis about gravel roads condition development.

The developed measurement method is unique because it measures and reports about quality of gravel roads in real time, which opens many opportunities to improve condition management and quality control of gravel roads. The real time quality control also allows quick information to road users if there is for example some kind of security risk.

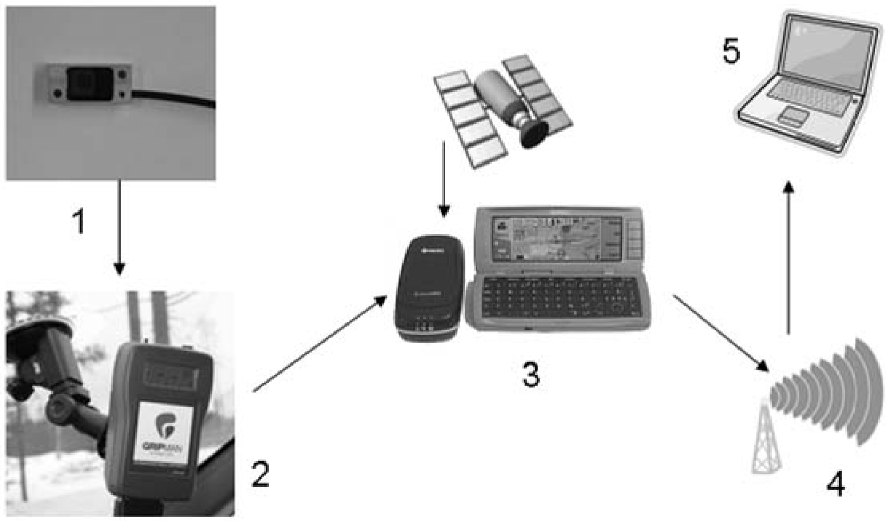


Figure 8. Parts which are included in the automatic measurement system: 1. The acceleration sensor which is placed on the rear right suspension arm of the vehicle. 2. The terminal box which process and sends via wireless the measurement information to mobile phone. 3. Information management software which includes GPS readings. 4. The measurement results are sent to the server. 5. User can review measurement information on web.

Figure 9. Classification of road condition is shown in real time on mobile phone screen. The decimal number on top of left corner shows road condition.

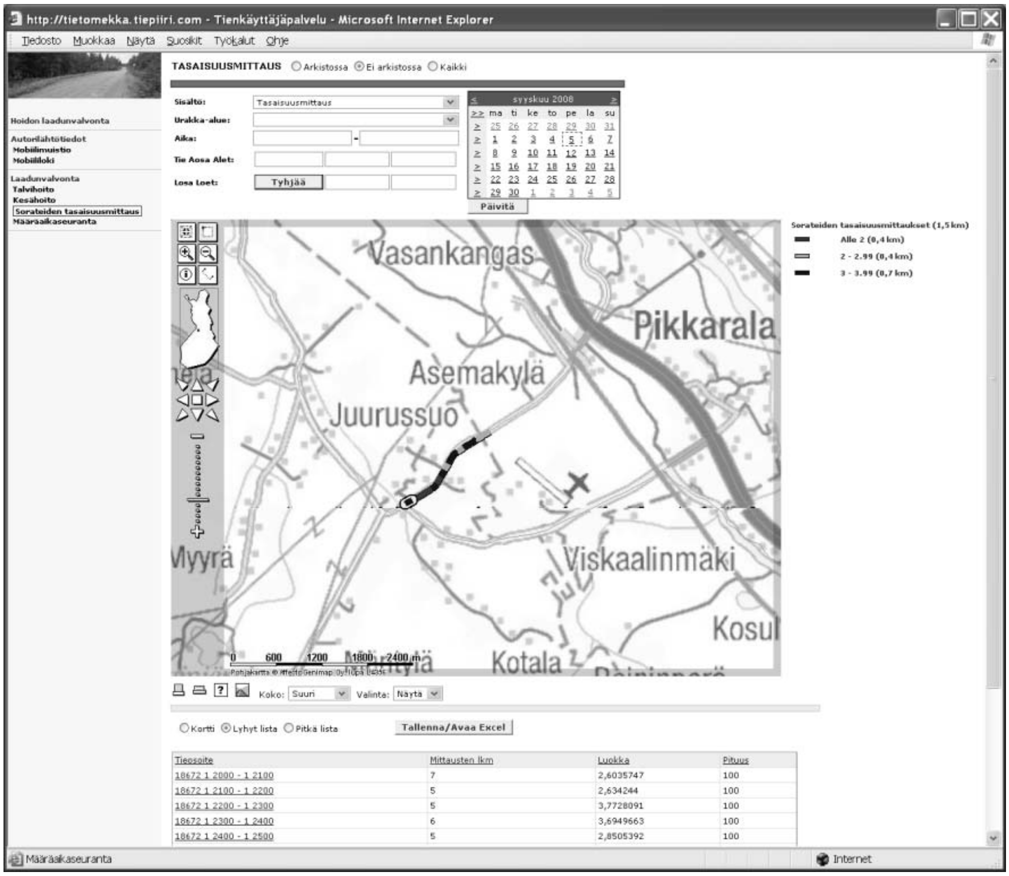


Figure 10. The measurement details are shown on web page. The map color describes the road condition.

References

[1] ISO 2631–1. (1997). Mechanical vibration and shock – Evalation of human exposure to whole-body vibration- part 1: General requirements.

[2] Lampinen, Anssi. (1998). Controlled management of gravel road, Condition rating and measuring method of gravel road. Al-Engineering Company. Research report . (in Finnish)

[3] Lampinen, Anssi. (2001). Controlled management of gravel road, Condition rating and measuring method of gravel road. Al-Engineering Company. Research report. (in Finnish)

[4] Marjanen, Ykä. (2002). Analysis of whole-body vibration on heavy vehicles. Department of Mechanical Engineering; University of Oulu, Final thesis. (in Finnish)

[5] Tietomekka company website <http://www.tietomekka.fi> (Dec. 9.2008)

Impact of Fast Automated Tracking of Construction Components on Labor Productivity

David Grau[1], Carlos H. Caldas[2], Carl T. Haas[3], Paul M. Goodrum[4], and Jie Gong[5]

[1] Assistant Professor, Department of Civil, Construction, and Environmental Engineering, The University of Alabama, Box 870205, Tuscaloosa, AL 35487-0205, Tel. (205) 348-6167, FAX (205) 348-0783, dgrau@eng.ua.edu
[2] Assistant Professor, Department of Civil, Architectural, and Environmental Engineering, The University of Texas at Austin, 1 University Station C1752, Austin, TX 78712-0273, Tel. (512) 471-6014, FAX (512) 471-3191, caldas@mail.utexas.edu
[3] Professor, Department of Civil Engineering, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, N2L 3G1, Canada, Tel. +1 (519) 888-4567, FAX +1 (519) 888-4300, chaas@civmail.uwaterloo.ca
[4]Associate Professor, Department of Civil Engineering, 151C Raymond Building, University of Kentucky, Lexington, KY 40506-0281, Tel. (859) 257-5416, FAX (859) 257-4404, pgoodrum@engr.uky.edu
[5] Ph.D. Candidate, Department of Civil, Architectural, and Environmental Engineering, The University of Texas at Austin, 1 University Station C1752, Austin, TX 78712-0273, Tel. (512) 471-8417, FAX (512) 471-3191, jie.gong@mail.utexas.edu

Abstract

Even though the fundamental importance of construction components in any given project, industry practices still solely rely on the human ability to individually control thousands of these components on the field. This lack of automation frequently results in critical errors that negatively affect project cost and schedule. Recently, though, the undemonstrated notion that materials tracking processes can highly benefit from the implementation of information technologies has been gaining wide industry acceptance. This paper presents the results of a massive study on a large industrial site that aimed at quantifying the impact associated with automating materials tracking processes on craft labor performance. For this purpose, field records from manual and automated tracking processes were collected during the trial. Then, the influence of the automated tracking process on construction performance was determined by considering the manual approach as the baseline for comparison. The results indicate that information technologies can significantly enhance craft labor productivity.

Motivation

Advances in technology have many benefits and among the most often cited are improved quality and productivity. The evolution within the communications industry can be seen as one example of how innovation has been embraced and leveraged on a continuing basis. Arguably, the construction industry lags in this regard and underutilizes advances in technology. However, the opportunity to improve construction productivity exists, and there is evidence that sectors of the construction industry have experienced long term productivity growth as a result.

Economic research has shown that technology tends to have a greater impact on labor productivity versus factor productivity measures. For example, investment in new equipment technology may improve an organization's labor productivity, but their factor productivity may actually decline if the relative increase in the cost of the equipment outweighs the relative savings in labor costs and gains in output. While there is evidence that improvements in equipment and material technology have a positive impact on construction productivity, the impact of information technology has remained largely undocumented.

Industry practitioners have recently identified site materials tracking as a process highly susceptible of being improved by new information technologies (Vorster and Lucko 2002). According to a wide survey among a large number of practitioners, new technology devices offer site materials management the greatest potential for improvement and the greatest potential for positive economic impact on engineering construction projects. Moreover, this synergistic integration also opens the possibility of doing more work with fewer people and reducing the number of safety incidents.

Even though the opportunity to advance material handling practices exists throughout the industry, industrial project sites offer the greatest potential for improvement. On these sites, the challenges that their materials handling practices need to satisfy steadily increase over time. As project demands and regulations increase over the years so does the number of engineered components. Actually, materials managers need to control thousands of unique components over several years and throughout areas covering tens or hundreds of acres. During their storage, a large percentage of components is frequently moved making their proper control unattainable and negatively impacting project performance. Projects with remote-site locations are a prominent example. That a component required in the critical path is lost and needs to be re-ordered to a distant manufacturer virtually guarantees that the component will not be available for installation when required. This kind of materials management oversight increases total installed costs and can eventually delay a whole project.

This imbalance between the magnitude of the tracking effort and the manual approach to it is reflected in the historically low efficiency of manual tracking processes. It has been consistently observed that workers waste up to a third of their working time in search of the items that they need or waiting idle for their availability while others are searching for them (Rojas and Aramvareekul 2003; Caldas et al. 2006). Anecdotal data also show that the amount of time needed to search for a particular component can randomly range from a few minutes to several hours (Grau 2008). In addition, an effective materials management system can not only prevent loosing components but also minimize the surplus of re-ordered pieces at the end of the project (CII 1999).

That there is not factual data demonstrating the potential benefits associated with the implementation of advanced tracking technologies seems to reasonably explain the lack of their actual adoption by the construction industry. Even if the industry assumes this advantage, the reality is that how these technologies actually impact craft labor performance and installation processes has not been previously studied. Overall, the questionable offset of this standing barrier is preventing the deployment of technology-supported tracking processes by the capital investment industry.

Background

A variety of technology-based approaches have been recently proposed to improve the monitoring of construction materials. These research efforts have primarily used Radio Frequency Identification (RFID), Global Positioning System (GPS), and several other sensing technologies to simplify and automate the tasks of identifying and locating construction components. Identification-based efforts were found to increase supply chain visibility when RFID was combined with barcodes (Navon and Berkovich 2006). This RFID technology also facilitated both the tracking of construction tools and the storing of their operational data in building environments (Goodrum at al. 2006). Other research has analyzed the potential of combining RFID technology with localization algorithms to locate construction components (Song et al. 2006; Grau and Caldas 2008).

The GPS positioning technology has also been placed on top of hard hats to infer labor activities from workers' positions (Sacks et al. 2003). GPS was also implemented on an industrial project to record the present time positions of pipe spools and was found to substantially reduce locating times (Caldas et al. 2006).

In spite of these successful applications of sensing technologies, the impact of automated tracking methods on project performance remains undetermined. Assessing and quantifying the benefits from automating site tracking practices can drive their actual implementation on the job sites.

Research Questions

This study is defined by two research questions:

1. Materials handling practices can be automated to effectively monitoring the status of a large number of components under challenging type of conditions, such as those of construction job sites.
2. The automation of materials tracking practices has a positive economic impact on project performance.

In order to satisfactory answer both questions; a research approach based on a trial run that compared the performances between manual and automated materials tracking processes was developed.

The Project Site

The purpose of this study is to estimate the impact of materials tracking technologies on craft labor productivity, labor efficiency, and percentage of missing components. Using a manual tracking process as a baseline, a massive trial run determined the productivity impact associated with the automated monitoring of steel components in a $750 million power plant project site ruled by open shop and direct hire policies. The project site had two almost identical boiler-support steel structures. Both steel structures were erected in equal sequences of installation with a time gap that oscillated between one and two weeks only.

In terms of materials, the project site was divided into two nearby areas: the lay down yard and the staging area. The lay down yard extended over an area of 25 acres and stored the structural steel elements from the moment they were received until the moment they were retrieved for their installation at the staging area. At the staging area, the components were prepared for their erection around the boiler structures in a very limited amount of space.

The Materials Tracking Process

Structural steel components were stored at the lay down yard directly from the trucks that delivered them. Adjacent grids partitioned the lay down yard area for location purposes. Each grid extended over an area of 30x15m2 and could be identified by a coded metallic post at its center. Each component was identified by its unique piece mark, which was engraved in a tiny steel plate that had been previously welded to the component. In order to make the identification of a component more evident to craft workers, its corresponding piece mark was written on the component at its delivery. Then, craft workers manually recorded the piece marks of the components and their corresponding grid locations for inventory purposes. Later, this information was typed into a project management system. If a steel element was moved to a different grid, crews were expected to record the new grid location data with the aim of updating it into the management system.

Based on the stored data, foremen submitted a materials withdrawal request (MWR) with a list of the components requested for installation. In order to minimize the presence of components in the space-constrained staging area, MWRs were submitted the same day or the day before the installation of their listed components. Based on the requested components and their grid-recorded locations, craft workers searched for them. If a component could not be located, workers utilized additional data from its detailed drawings to identify it. Once found, craft workers tagged the component with a colored tape. Different tape colors were assigned to different boiler installation sequences. Hence, tagging allowed erection crews to immediately identify and load the components to be retrieved onto flat bed trucks. That the hourly cost of craft workers was half of the hourly cost of erection crews and that this later crews operated costly pieces of equipment explains why management preferred craft workers for the time-consuming search of lay down yard components. Once beyond the lay down yard, components were not tracked.

At the staging area, components were stored nearby their rigging position in order to minimize their movement. Once lifted and placed into position, other crews bolted, plumbed, aligned, painted, and inspected each steel component.

Opportunities for Positively Affecting Labor Performance

Even though the materials management system was the most sophisticated the authors have observed over the years, advanced technology devices opened the door for its improvement. The opportunities for improvement are summarized below:

- *Reduction of locating times at the lay down yard.* The manual tracking process prevented efficient and reliable searches. Frequently, components were in a different grid than those indicated by records. In these cases, workers would need to spend unnecessary amounts of time searching for the components that they could not find. Even in the instances of correctly inventoried components, the 450m^2 grid-size made searches inefficient.
- *Percentage of not-immediately-found components.* Craft workers had to frequently re-search for components that had not been found using the available information. These random searches required four or five workers on average and could last an entire day, further extending location times.
- *Reliable support for installation processes.* The inability to immediately supply the staging area with steel components for their installation decreased erection productivity.

- *Monitoring of components in the installation area*. The status of components pending for installation at the staging area was always unknown. This lack of information further jeopardized the installation process by creating confusion among erection crews. Thus, erection crews were observed to spent considerable amounts of time figuring out the whereabouts of the components planned for installation.

Automated Materials Tracking Technology

The automated materials tracking process is based on the combination of localization algorithms, and GPS and RFID technologies. Initially, RFID tags were attached to the components to be tracked. Then, the positioning and identification sensors were hung together from a piece of roving equipment. While roving throughout the jobsite, the GPS receiver determined its own position and the RFID reader identified the signals from the emitting components around this position. Then, localization algorithms estimated the location of the tagged components by processing the time series of location and identification data. On average, the resulting accuracy of these estimated locations was better than ten feet. Maps representing the position of the estimated location of the components could be generated to support field searches. Grau and Caldas (2009) provide a comprehensive description of this automated tracking technology.

Trial Design

The trial design was based on the existence of two identical boiler (hereafter named A and B) units with equivalent sequences of installation. The authors utilized these matching conditions coupled with the fact that each boiler had its independent foreman and crews to the advantage of this study. These ideal conditions enabled the direct comparison of unbiased productivity measures. While 400 Boiler B components were tagged and abided by the automated tracking process, their equivalent Boiler A components abided by the manual tracking process. Then, site-wide records from both sets of components were collected in tandem from their storage at the lay down yard until the moment they were lifted for installation. Both sets of collected records were used to determine the impact of the automated tracking process on project performance.

On the lay down yard, the authors collected the amount of craft worker time to inventory, locate, and flag the components. While MWRs for Boiler A were supported with a list of component identification and grid location information, MWRs for Boiler B components were supported with user-friendly maps representing the estimated locations of the required components. On the staging area, the contractor continuously collected engineering data, project controls data, and foreman delay surveys for each boiler unit. While the traditional tracking process did not monitor the components in the staging area, the authors provided the Boiler B foreman with an updated list of the components available for installation and an equivalent map with their corresponding positions.

Performance Metrics

A set of three metrics was defined to determine the impact of the automated tracking approach on project performance. These metrics were:

- *Craft labor productivity at the lay down yard*. The average time to inventory, locate, and flag the steel components.
- *Percentage of components not-immediately-found at the lay down yard*. The percentage of components that could not be located in an immediate manner with positioning and identification data and required extended search efforts.
- *Steel erection productivity*. The weight of erected components per work-hour —inclusive of the hours required for unloading, storing, identifying, locating, and erecting the steel at the staging area. The work hours required for the activities beyond the moment the materials had been lifted —such as bolting, plumbing, torquing, painting, or inspecting— were not considered due to their independence with respect to the materials management process.

Trial Results

The trial evidenced a strong potential for improving project performance by automating site tracking processes. On the lay down yard, the automated tracking process minimized both the average labor time spent per component and the number of not-immediately-found components. In detail, the average labor time spent per steel component on the yard was reduced in a ratio of eight to one (See Table 1), a difference that proved to be statistically significant. This reduction was due to both the precise map estimated positions (versus the 450m2 accuracy of the grids) and the minimized number of not-immediately-found components. Indeed, the percentage of components that required extended searches decreased from 9.53% to 0.54% (See Table 2). This is equivalent to say that for each twenty manually-tracked steel components that could not be immediately found by craft workers there was only one instance of automatically-tracked component. It was observed that craft workers were confident of the precise map locations to the extent that in the several instances of steel components that could not be rapidly found, craft workers directly reported the issue without extending their search to adjacent areas. However, craft workers had to inevitably perform these extended searches when components could not be found based on their corresponding grid locations.

Table 1. Average craft labor time spent per lay down yard component

Tracking Method	Minutes
Traditional	36.80
Automated	4.56

Table 2. Percentage of components with extended search times

Tracking Method	Percentage of Not-Immediately-Found Components
Traditional	9.52
Automated	0.54

As a consequence of these differences, the average labor time per steel components was 36.80 minutes for the manual process and only 4.56 for the automated process (See Table 3). For the manual process, craft workers devoted 8.23 for inventorying, finding, and flagging each component, while only 2.93 minutes were required for the automated process. Complementarily, craft workers spent an average of 28.57 minutes per component in extended searches of non-immediately-found components for the manual process, while they only had to spend an average of 1.63 minutes for the automated process. These extended search times can be directly regarded as waste from a productivity perspective, being their values indicative of the gross potential for improvement of each type of tracking process.

Table 3. Average craft labor times per component

Traditional Tracking Method

Activity	Minutes
Mark Component	1.09
Record Grid Position	0.41
Locate and Flag	6.73
Extended Searches	28.57
Average Time	36.80

Automated Tracking Method

Activity	Minutes
Collect Data	0.73
Locate and Flag	2.20
Extended Searches	1.63
Average Time	4.56

On the staging area, the erection productivity of tagged components incremented in 4.2% when compared to the components manually tracked (See Table 4). Since the contractor credited 50% of the installation right after the erection of steel elements ─without considering the remaining activities of aligning, plumbing, bolting, and painting, this productivity increment is equivalent to a 2.1% increment in the overall installation productivity. However, the authors also analyzed the impact of the automated tracking process in the subset of components that were lighter than 500Kg, such as small beams and hollow sections. As it could have been expected, the automation effect was higher on this more prone-to-loose subset of components, resulting in a 13.4% increment of erection efficiency.

Table 4. Erection productivity improvement

Components	Improved Productivity (%)
All	4.2
Less than 500 Kg	13.4

Cost Effectiveness at the Project Site

For a tracking technology to be feasible, it does not only need to result in increased productivity but also needs to be cost-effective. A benefit-to-cost analysis was performed to determine whether the benefits surpassed the upfront technology costs. Assuming that the 9,670 steel components had been tagged, almost a 2:1 benefit-to-cost ratio would have been achieved (See Table 5). In the pessimistic case that the fifteen-dollar tags could be re-used only once ─these tags have a guaranteed battery life of six years, a close to 4:1 benefit-to-cost ratio would result ─since benefits would double. Subsequent re-uses would further increase this highly positive benefit-to-cost ratio. In addition to these economic benefits, the installation schedule at the trial site could have been reduced in almost six days per boiler unit as a result of a more efficient erection process.

Table 5. Cost benefit of tagging all the steel components

Concept	Amount
Lay Down Yard Savings	$103,920
Staging Area Savings	$169,337
Technology Costs	($151,750)
Benefits	$121,507

There are two important remarks that the reader needs to be aware of. First, the contractor had implemented a very sophisticated materials management process, highly above the industry average in terms of efficiency. Over the years, the authors of this research have been able to observe that the degree of sophistication of site materials management practices varies widely. While some companies do not have any specified procedure to control site components, others have developed exhaustive site materials management processes based on project management information systems. Based on these observations it can be stated that the materials management process taken as the baseline for this study is much more sophisticated than an average industry process. Second, heavy steel items difficult to misplace or loose were the objective of the tracking process. Other objects, such as pipe spools or valves, can be more easily misplaced or lost. Overall, implementing a process for automatically tracking more prone-to-loose type of components on average type of projects can result in positive impacts far beyond those obtained in this study.

Conclusions

This study assessed and quantified the impact associated with the automation of site materials tracking processes on project performance. For this purpose, a massive field trial on a large industrial site was performed on the premises of tracking structural steel elements. A sophisticated traditional tracking process was the baseline for the improvements determined by this study.

The tracking approach surpassed the manual approach in every quantified aspect. The average labor time spent per tracked component was reduced in a ratio of eight to one. This significant labor time reduction resulted not only in more efficient and predictable searches, but also increased the moral of the workers. In reality, the number of temporarily missing components was almost eradicated. This facilitated the continuous retrieval and installation of these components. As a consequence, the erection productivity was significantly increased.

These savings were demonstrated to potentially result in significant project benefits if the automated tracking methodology had been widely applied. Important cost savings and schedule reductions could have been attained to justify the implementation of these tracking technologies. These tangible benefits are critical to drive the implementation of materials tracking technologies on construction job sites.

Future efforts should address the implementation of these tracking technologies from manufacturing to installation. Indeed, they could record and store component information regarding design, procurement, fabrication, delivery, transportation, receiving, installation, and maintenance. The instant visibility of construction elements through their life-cycle could significantly improve the efficiency of the capital investment industry. Such a thorough-going advancement stands to become another driving force for opening the doors to these and other technology solutions within the industry.

Acknowledgements

The authors would like to acknowledge the Construction Industry Institute and FIATECH for their financial support. They would also like to recognize the support received from Bechtel and Identec.

References

[1] Caldas, C. H., Grau, D., and Haas, C. T. (2006). "Using Global Positioning System to Improve Materials-Locating Processes on Industrial Projects." Journal of Construction Engineering and Management, 132(7), 741-749.

[2] CII (1999). "Procurement and Materials Management: A Guide to Effective Project Execution." The Construction Industry Institute, Implementation Resource 7-3.

[3] Goodrum, P.M., McLaren, M.A., and Durfee, A. (2006). "The Application of Active Radio Frequency Identification Technology for Tool Tracking on Construction Job Sites." Automation in Construction, 15(3), 292-302.

[4] Grau Torrent, D. (2008). "Development of a Methodology for Automating the Localization and Identification of Engineered Components and Assessment of its Impact on Construction Craft Productivity." Ph.D. Thesis, The University of Texas at Austin, Austin, Texas.

[5] Grau, D., and Caldas, C.H., (2009). "Methodology for Automating the Identification and Localization of Construction Components on Industrial Projects." Journal of Computing in Civil Engineering, 23 (1), 3-13.

[6] Navon, R., and Berkovich, O. (2006). "An Automated Model for Materials Management and Control." Construction Management and Economics, 24(6), 635-646.

[7] Rojas, E. M., and Aramvareekul, P. (2003). "Labor Productivity Drivers and Opportunities in The Construction Industry." Journal of Management in Engineering, 19(2), 78-82.

[8] Sacks, R., Navon, R., and Goldschmidt, E., (2003). "Building Project Model Support for Automated Labor Monitoring." Journal of Computing in Civil Engineering, 17 (1), 19-27.

[9] Song, J., Haas, C.T., and Caldas, C.H. (2006). "Tracking the Location of Materials on Construction Job Sites." Journal of Construction Engineering and Management, 132(9), 911-918.

[10] Vorster, M. C., and Lucko, G. (2002). "Construction Technology Needs Assessment Update." The Construction Industry Institute, Research Report 173-11.

Evolutionary Fuzzy Hybrid Neural Network for Conceptual Cost Estimates in Construction Projects

Min-Yuan Cheng[1], Hsing-Chih Tsai[2], and Erick Sudjono[3]

[1, 2, 3] Department of Construction Engineering, National Taiwan University of Science and Technology; P.O. Box 90-130, Taipei, Taiwan. FAX (8862) 2730-1074.
[1] Professor; PH (8862) 2733-6596; e-mail: myc@mail.ntust.edu.tw
2 Project-appointed Assistant Professor; PH (8862) 2730-1277; e-mail: tsaihsingchih@gmail.com
3 Master; PH (8862) 2737-6663; e-mail: erick_jono@yahoo.com

Abstract

Conceptual cost estimates are important to project feasibility studies, even the final project success. The estimates provide significant information for project evaluations, engineering designs, cost budgeting and cost management. This study proposes an artificial intelligence approach, the evolutionary fuzzy hybrid neural network (EFHNN), to improve precision of conceptual cost estimates. The approach incorporates neural networks (NN) and high order neural networks (HONN) into a hybrid neural network (HNN). The HNN operates with the alternative of linear and non-linear neuron layer connectors. Besides, fuzzy logic (FL) is employed for handling uncertainties, the approach therefore evolve into a fuzzy hybrid neural network (FHNN). For FHNN optimization, the genetic algorithm is used for both FL and HNN, consequently the approach is named as EFHNN. In practical case studies, two estimates including overall and category cost estimates are provided and compared. Results show that the proposed conceptual cost estimates can be deployed as accurate cost estimators during early stages of construction projects. Moreover, considering linear and non-linear neuron layer connectors in EFHNN surpasses models with singular linear deployment of NN.

Keywords： Construction Cost, Conceptual Estimates, Genetic Algorithm, Fuzzy Logic, Neural Network, High Order Neural Network, Hybrid Neural Network

1. Introduction

Cost estimates are fundamental to all project-related engineering and greatly influence planning, design, bidding, cost management/budgeting and even construction management. Such estimates allow owners and planners to evaluate project feasibility and control costs effectively in detailed project design work. Due to the limited availability of information during the early stages of a project, construction managers typically leverage their knowledge, experience and estimators to estimate project costs, i.e., they usually rely on their intuition. Researchers have worked to develop cost estimators that maximize the practical value of limited information available in order to improve cost estimate accuracy and reliability, which should improve the suitability of resultant designs and subsequent project execution work.

Artificial intelligence approaches are applicable to cost estimating problems related to expert systems, case-based reasoning (CBR), neural networks (NN), fuzzy logic (FL), genetic algorithms (GA), and derivatives of such. Many research studies have been done in this area. For instance, Serpell (2004) proposed a model of this problem based on existing knowledge and demonstrated how the model was used to develop a knowledge-based assessment system. An et al. (2007) developed a case-based reasoning model, where an analytic hierarchy process was employed to incorporate experience. NN represents the most frequently applied approach in this type of application. Wilmot and Mei (2005) developed an NN model to estimate highway construction cost escalation over time.

Furthermore, hybrid models have also been developed to estimate construction costs. Hegazy and Ayed (1998) used NN to develop a parametric cost estimating model for highway projects, with optimal NN weightings optimized by GA. Yu et al. (2006) proposed a web-based intelligent cost estimator that incorporated a neurofuzzy system.

In past research work, NN, GA, and FL have been employed for their powerful abilities to estimate construction costs. They are widely applied as well to other topics and even to fields unrelated to the construction industry.

High order neural networks (HONN) usually introduced a nonlinear equation to a specified layer. It allowed the networks to capture high order correlations easily and attained the nonlinear mapping effectively. As HONN uses high order correlations, it may perform better than linear NN (Zurada, 1992). HONN not only allowed a fuller degree of adaptability in the form of the nonlinear mapping than linear model, but has a structure that can make it easier to determine how the network inputs come to be mapped into the network outputs (Abdelbar, Tagliarini, 1996).

In previous studies (Cheng et al. 2008a; Cheng et al. 2008b), the authors had contributed on a GA-optimized Neural-fuzzy model for studying various engineering problems. This study incorporates the linear neural networks (NN) and high order neural networks (HONN) as a hybrid neural network (HNN). Each HNN layer connector is dominated by an alternative of selecting a linear or high order layer connector. The HNN model evolved into a fuzzy hybrid neural network (FHNN) model with participation of fuzzy logic. Within the proposed evolutionary fuzzy hybrid neural network (EFHNN) model, GA optimized both of FL's membership functions and HNN's connection types, topology, and coefficients, etc. This study further applied proposed EFHNN for the uses of conceptual cost estimators. Two kinds of estimates - overall estimates (total cost estimates) and category estimates (estimating respectively with engineering categories) – were provided at the planning or preliminary design stage.

2. Evolutionary Fuzzy Hybrid Neural Network (EFHNN)

The proposed EFHNN incorporates four artificial intelligence approaches which are neural network (NN), high order neural network (HONN), fuzzy logic (FL), and genetic algorithm (GA) (see Fig. 1). Of which, NN and HONN are composed for the inference engine, i.e. the proposed hybrid neural network (HNN); FL dominates both fuzzifier and defuzzifier layers; and GA optimizes the HNN and FL. According to the definitions of "neuro with fuzzy input-output" (Hayashi et al. 1998), this study proposes a fuzzy hybrid neural network (FHNN) which is structured of HNN with both fuzzy inputs and fuzzy outputs (see Fig.2). Each NN connection can select a linear or high order NN connector. Sequentially, the FHNN is optimized through GA adaption process (see Fig. 3). The process simultaneously searches for the optimum FL's membership functions, defuzzification coefficients, HNN topologies, and HNN parameters (including linear/high order connection types) using GA. In the process, P(t) denotes a population at generation t, PO(t) is an offspring population at generation t, and PM(t) indicates a mutation population at generation t. Details of FL and HNN and GA are described in the following sections.

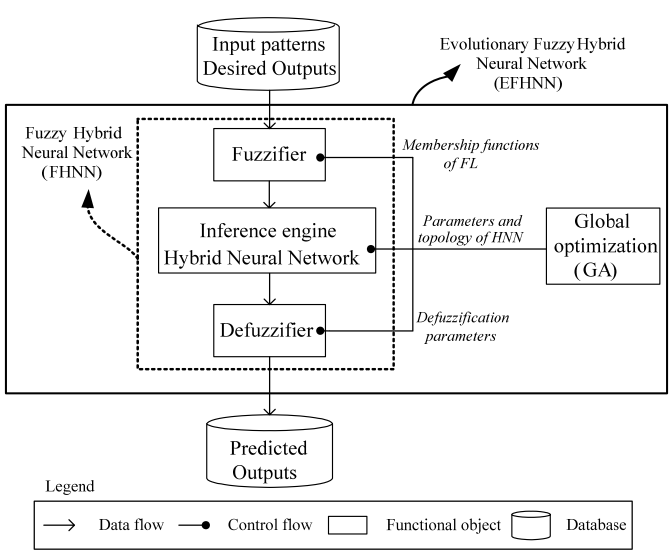


Fig. 1. EFHNN Architecture

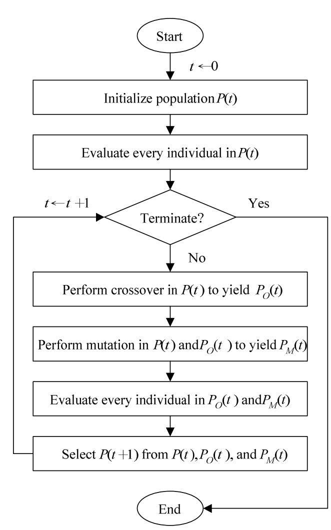


Fig. 3. EFHNN Adaption Process

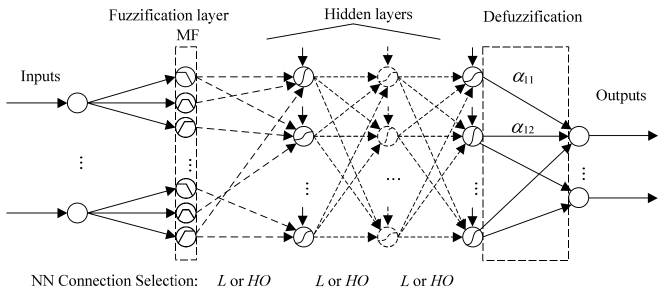


Fig. 2. FHNN with FL and HNN

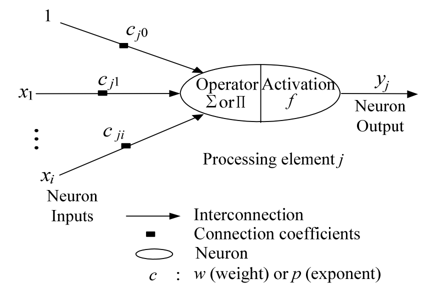


Fig. 4. A HNN Neuron

2.1 Proposed Hybrid Neural Network

Originally, "hybrid" means anything derived from heterogeneous sources, or composed of elements of different or incongruous kinds. For the proposed hybrid neural network (HNN), "hybrid" is used for representing the combination of traditional neural network and high order neural network. The high order neural network that this paper uses was proposed by HONEST model (Abdelbar and Tagliarini, 1996) which was structured of three layers with a high order connection and a linear connection between 1st – 2nd layers and 2nd – 3rd layers respectively . This study extends uses of high order connection for all connection alternatives, i.e. all layer connections can switch to linear or high order type (see Fig.2). An HNN neuron is dominated by an alternative of following equations:

Linear connection: $$y_j = f\left(\sum w_{ji} x_i + b_{j0} \times 1\right) \quad (1)$$

High order connection: $$y_j = f\left(\prod x_i^{p_{ji}} \times 1^{b_{j0}}\right) \quad (2)$$

Activation function: $$f(x) = \frac{1}{1+e^{-ax}} \quad (3)$$

where yj is a HNN neuron output calculated by neuron inputs xi. cji represents a coefficient of a interconnection, which can be linear or high order format related to a weight wji or exponent pji respectively (see Fig. 4). An activation function f uses sigmoid function with a slope coefficient of a. Therefore, each layer connection has an attached connection type to represent thereof operation selection (see Fig. 2). All the HNN parameters will then be optimized by GA evolution. As above mentions, a HNN with 2 layers can select a linear layer connection (L) or a high order connection (HO); with 3 layers number of HNN models according to connection types, four possible scenarios of L-L, L-HO, HO-L, and HO-HO exist. If adopting a maximum number of HNN layers as N, i.e. the final HNN model can be a HNN with layer not beyond N, there are 21, 22,…, 2N-1 HNN model candidates respectively related to HNN with 2, 3,…, N layers. Summarily, there are 2N-2 HNN model candidates. Of which, only N-1 models select all L connections, all the remainders are categorized into high order neural networks in this study. The proposed HNN includes all the linear and high order neural networks according connection type selections.

2.2 Fuzzy Logic Facilities

In Fig. 2, the HNN is enclosed by a fuzzification layer and a defuzzification layer. All of these compose a fuzzy hybrid neural network. In the defuzzification layer, an input firstly transfers into several membership grades by membership functions (MF). In this study, a complete MF set which uses trapezoidal MF is adopted. A general way to describe the shapes of MF is to depict summit positions (smi) and widths (wdi) of MF (Ishigami et al., 1995; Hayashi et al., 1998). An input can be transferred into several membership grades with the membership functions. Originally, the membership function inputs of are bounded between the range of layer inputs and the membership function inputs are usually set within [0,1]. However, owing to the adopted equation (2), while one of the membership function outputs has zero value, the related HNN neurons will output zero values through the sigma-pi Π operator. To prevent such, this study modified the original MF to the output range of [0.0001,1] (see Fig. 5). Following the aforementioned descriptions, all the membership functions are characteristic of values of sm and wd. Besides, in defuzzification layer (see Fig. 2), this study adopts weighted average formula:

$$y_i = \psi(x) = \frac{\sum \alpha_{ji} x_i}{\sum \alpha_{ji}} \qquad (4)$$

where ψ is a defuzzification function; α represent defuzzification weights; x denotes the eventual outputs of HNN; and y are the final FHNN outputs. Consequently all the sm, wd, and α will then be dominated by GA evolution.

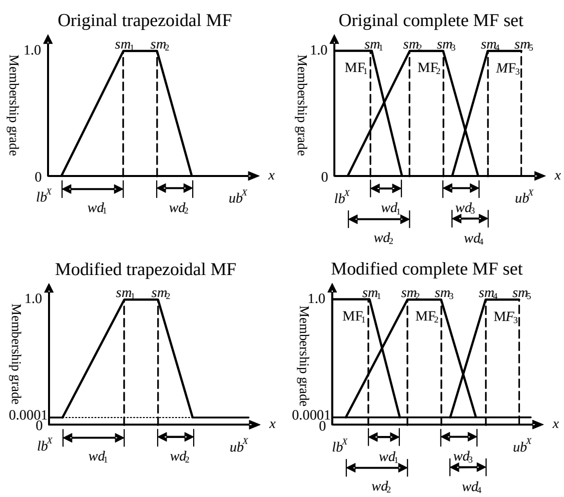


Fig. 5. Membership Function Example

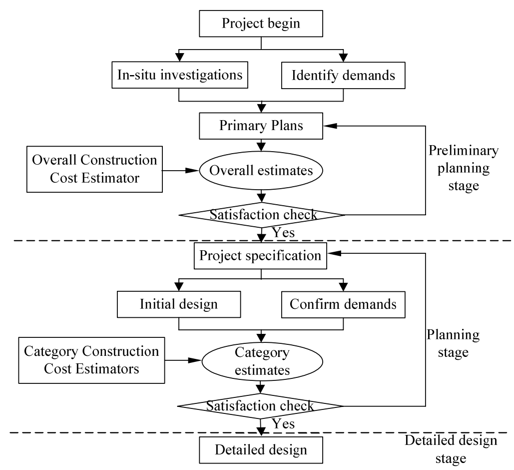


Fig. 6. Cost estimators during project planning stage

2.3 Genetic Algorithm Facilities

To apply GA for problem optimization, one should identify all essential parameters to determine the length of chromosome. A chromosome (an individual) in this study is to represent a FHNN with parameters of HNN and FL. HNN's Parameters have interconnection coefficients c (w and p), connection types (CT: L or HO), slope coefficient of activation function a (1~6), and network topology (number of layers and number of layer neurons). FL's parameters include MF summit points (sm), MF widths (wd), and defuzzification weights α. It deservers to be mentioned is that an interconnection coefficient c can be used for the alternative of w or p. However, w and p perform totally different, they should be recorded in different sub-string. Therefore the aforementioned c should be a combination of w and p. As the chromosome of an individual is identified, the FHNN can be optimized through the adaption process with crossover, mutation, and selection mechanisms (see Fig. 3). Each model result is evaluated with the root mean square error (RMSE).

3. Conceptual Cost Estimators

In the planning stage of projects (i.e. initial design has not yet been produced), the overall estimator is identified by six quantitative factors and four qualitative ones (see Cheng 2008c). These factors will be treated as EFHNN inputs.

Once a project design has been drafted, category cost estimators can be employed to calculate engineering cost by categories. Comparing with the overall estimate, the summation of all category estimates is another alternative. Therefore, the category estimators will be more applicable and useful to the overall estimator for project management. One category estimate will be evaluated for each engineering category according to particular factors. There were seven types of engineering work generalized for category construction cost estimates (see Cheng 2008c).

The range of construction project data spans the years 1997 through 2001. Construction cost range was assigned between NTD40,179 and NTD98,285 per square meter. All 28 projects were designed using reinforced concrete in main structural members. All but five (23) cases were used as training cases, with the remainder (5) used for testing the approach in this paper. Ten inputs were set as the overall construction cost estimator and one output served as the overall estimate of total unit cost (i.e., construction cost per square meter). Seven category estimates (respective outputs, i.e. unit cost by categories) are calculated by engineering category (see Cheng 2008c), where 4 inputs for temporary construction; 7 for geotechnical construction; 8 for structural construction; 9 for interior decoration; 8 for electromechanical infrastructure; 5 for miscellaneous construction; and 4 for indirect construction. Construction costs used as training targets reflect Taiwan's published price index for calendar year 2001. Therefore, the proposed estimators are capable of dealing with unit price fluctuation of the wok and material items in the market. These estimators were developed to meet the goal of assisting construction project planning and design through the use of evaluated cost estimates. In Fig. 6, an overall construction cost estimator is used in the preliminary planning stage, when detailed project plans have not yet been drafted. Preliminary plans can be drafted with in-situ investigations and identified demands, after which the generated overall cost estimate can be used to check the relevance and accuracy of those plans. Initial design will be done next in the planning stage, and then demands and designs will be checked against category estimates. Detailed planning and design can be executed once all data and estimates meet project management needs. These conceptual estimates significantly influence project construction and management.

4. Results and Comparisons

As noted above, this study developed two kinds of estimators and used 23 training cases and 5 testing cases. The capabilities of EFHNN were employed in these estimates. However, the EFHNN is time-consuming due in large part to its use of GA. Therefore, experiments should be run to set parameters to a practicable range (see Table 1). All the results will be compared with the results of previous model (evolutionary fuzzy neural inference model, EFNIM, which did not employ high order neural network and changes in FL and GA).

Table 1 Parameter Settings for EFHNN

Parameters	Values
No. of input neurons	Number of influenced factors
No. of output neurons	1
No. of maximum hidden layers	5
No. of maximum neurons in each layer	5
Selected activation function	Logistic Sigmoid Function
Slope of activation function	1~6
Shape of membership function	Trapezoidal
Number of membership functions	5
Crossover rate	0.9
Mutation rate	0.025
Population size	50
Iteration set	5000

4.1 Overall construction cost estimator results

This estimates construction cost, obtained while the project is in the preliminary concept stage, i.e. without detailed engineering plans organized by category, has significant bearing on detailed planning. After a process of evolutionary training, achieved through applying the 23 training cases, five testing results are shown in Table 2.

Table 2. Testing Results of Overall Estimates

Case No.	Actual output (NTD/m2)	Desired output (NTD/m2)	Diff. (NTD/m2)
1	49697	61591	-11894
2	63763	56334	7429
3	51988	49139	2849
4	87454	84631	2823
5	63654	70843	-7189

Note: Diff. = Actual – Desired.

4.2 Category construction cost estimator results

Although the overall construction cost estimator has been developed, construction plans in each category must still be designed. Construction costs for engineering categories should be estimated to ensure costs are controlled effectively and facilitate project management. Although it is difficult to completely categorize construction work into types of engineering, such is essential in order to estimate category cost values and helpful in project planning and design. Table 3 not only shows estimation results, but also category cost ratios (take the structural category as an example). It is apparent that category cost ratios bear significantly on project planning and design. This result allows cost management to be effectively implemented into construction engineering categories.

Table 3. Testing Results of Category Estimates

Engineering categories	Case No.	Actual output (NTD/m2)	Desired output (NTD/m2)	Diff. (NTD/m2)	Ratio of category cost (%)
Structural construction	1	17398	18843	-1444	28.00
	2	16721	15795	926	28.15
	3	15725	15781	-55	30.67
	4	15726	14531	1195	20.99
	5	17416	17777	-360	26.54

Table 4. Result Comparisons of Overall and Category Estimates

Case No.	EFHNN prediction errors		EFNIM prediction errors	
	Error of overall estimates (%)	Error of total category estimates (%)	Error of overall estimates (%)	Error of total category estimates (%)
1	19.312	0.900	20.541	2.504
2	13.187	5.452	23.783	7.458
3	5.797	4.349	21.201	9.699
4	3.336	11.447	5.082	10.018
5	10.148	7.373	9.755	4.082
Avg.	**10.356**	**5.904**	**16.072**	**6.753**

4.3 Comparisons between EFHNN and EFNIM

In practice, overall estimates accurate to within 25% and category estimates accurate to within 15% using engineers' experience are typically considered acceptable. Estimators developed in this paper achieve a high level of precision for construction cost estimation during the early stages of a project (see Table 4). Estimating construction costs more precisely will help make designs more feasible and projects more efficient by enhancing project management. Moreover, the proposed EFHNN which employs both linear and non-linear layer connectors surpasses the previous version EFNIM which use the traditional NN connection only in the conceptual cost estimates (Cheng et al., 2008c).

5. Conclusions

This paper presents comprehensive descriptions of the proposed Evolutionary Fuzzy Hybrid Neural Network and thereof application of construction conceptual estimators. The EFHNN mechanism is a fusion of HNN, FL, and GA. HNN is composed of a traditional neural network (linear) and a high order neural network, FL uses a fuzzification and a defuzzification layers to sandwich the proposed HNN for a FHNN model, and GA is used to optimize FHNN's parameters for the proposed EFHNN. EFHNN innately different from various GA-FL-NN approaches, even if the previously proposed EFNIM, basing on layer connection types of HNN, modifications for FL's membership functions, and GA-optimized parameters. Therefore, the EFHNN address problems further with a huge amount of HNN models along with fuzzy concepts and GA optimization.

This study proposed two kinds of construction cost estimators. The overall construction cost estimator was established to estimate a total cost in the absence of categorized engineering plans. The category estimators, with additional data inputs, were established to evaluate engineering costs within categories. The advantages of proposed estimators:

1. An overall construction cost estimate can be provided during the preliminary project planning stage to facilitate project execution even when only a minimal amount of available data is available.
2. Category construction costs, categorized by engineering type, offer an alternative to overall estimates that provides results that are more reasonable and practicable.
3. Category estimators supply useful information on the relative ratios of engineering categories, which is essential for detailed construction cost management.
4. All estimates, come from EFHNN results, address problems with a newly developed HNN architecture with figure out an input-output mapping with both linear and nonlinear layer connections.
5. The EFHNN results for construction conceptual cost estimates surpass results from EFNIM which use the traditional NN connection only. It evidences that the HNN concept not only makes NN innately different but also performs well in EFHNN with both FL and GA.

This paper presents the application EFHNN to estimate construction costs during the early stage of construction projects in order to facilitate designers, owners and contractors for decision-making. Results show that EFHNN is relevant and applicable to construction management in Taiwan and may be implemented worldwide with modifications to account for specific regional/national factors.

References

[1] Abdelbar, A., Tagliarini, G., (1996). "HONEST: a new high order feedforward neural network." IEEE International Conference on Neural Networks - Conference Proceedings, 2, 1257-1262.

[2] An, S.H., Kim, G.H. and Kang, K.I. (2007). "A case-based reasoning cost estimating model using experience by analytic hierarchy process." Building and Environment, 42(7), 2573-2579.

[3] Cheng, M.Y., Tsai, H.C., Ko, C.H., and Chang, W.T., (2008a). "Evolutionary Fuzzy Neural Inference System for Decision Making in Geotechnical Engineering." ASCE Journal of Computing in Civil Engineering, 22(4), pp. 272-280.

[4] Cheng, M.Y., Tsai, H.C., and Liu, C.L., (2008b). "Artificial Intelligence Approaches to Achieve Strategic Control over Project Cash Flows." Automation in Construction. Accepted.

[5] Cheng M.Y., Tsai H.C., and Hsieh W.S., (2008c), “Web-based Conceptual Cost Estimates for Construction Projects Using Evolutionary Fuzzy Neural Inference Model.” Automation in Construction. Accepted.

[6] Hayashi, I., Umano, M., Maeda, T., Bastian, A. and Jain, L. C. (1998). “Acquisition of fuzzy knowledge by NN and GA.” IEEE International Conference on Knowledge-Based Intelligent Electronic Systems, 69-78.

[7] Hegazy, T. and Ayed, A. (1998). “Neural network model for parametric cost estimation of highway projects. Journal of Construction Engineering and Management.” 124(3), 210-218.

[8] Ishigami, H., Fukuda, T., Shibata, T., and Arai, F. (1995). “Structure optimization of fuzzy neural network by genetic algorithm.” Fuzzy Sets and Systems, 71(3), 257–264.

[9] Serpell, A.F. (2004). “Towards a knowledge-based assessment of conceptual cost estimates.” Building Research and Information, 32(2), 157-164.

[10] Wilmot, C.G. and Mei, B. (2005). “Neural network modeling of highway construction costs.” Journal of Construction Engineering and Management, 124(3), p 210-218.

[11] Yu, W.D., Lai, C.C. and Lee, W.L. (2006). “A WICE approach to real-time construction cost estimation.” Automation in Construction, 15(1), 12-19.

[12] Zurada, J.M., (1992). “Introduction to Artificial Neural Systems.” West Publishing Company, ST. Paul.

K-means clustering and Chaos Genetic Algorithm for Nonlinear Optimization

Cheng Min-Yuan and Huang Kuo-Yu

Dept. of Construction Engineering, National Taiwan University of Science and Technology, Taipei, 106, Taiwan, R.O.C

Abstract

To reduce the computational amount and improve estimation accuracy for nonlinear optimizations, a new algorithm, K-means clustering with Chaos Genetic Algorithm (KCGA) is proposed, in which initial population are generated by chaos mapping and refined by competition. Within each iteration of this approach, in addition to the evolution of genetic algorithm (GA), the K-means Clustering algorithm is applied to achieve faster convergence and lead to a quick evolution of the population.

The main purpose of the paper is to demonstrate how the GA optimizer can be improved by incorporating a hybridization strategy. Experimental studies revealed that the hybrid KCGA approach can produce much more accurate estimates of the true optimum points than the other two optimization procedures, the chaos genetic algorithm (CGA) and GA. Further, the proposed hybrid KCGA approach exhibits superior convergence characteristics when compared to other algorithms in this paper separately. On the whole, the new approach is demonstrated to be extremely effective and efficient at locating optimal solutions and verified by an empirical example from construction.

***Keywords**:* Chaos; optimization; K-means clustering; Genetic Algorithm

1. Introduction

In order to find a global or near-global optimal solution, the search by Genetic algorithm (GA) was a group base instead of the point-to-point search. The group, which contains several solution points, is named population and is represented by P(t) with t denoting the number of generations. The well known GA was introduced by Holland in 1970s as optimization approach. The main concept of this approach was derived from biological evolution in a competitive environment (Holland, 1975). Nowadays, many industrial applications have been developed with the aid of this tool (Hugget, Se´bastian & Nadeau, 1999).

At the meantime, GA are highly parallel randomly searching algorithms that imitate the life evolution as proposed in Darwinian survival of the fittest principle (Hibbert, 1993; Lavine and Moores, 1999). Critical genetic operations such as the encoding of the solution of optimizing problem, the designing of the fitting function according to its application, and the crossover and mutation for offspring, play important roles in GA (Holland, 1975; Zhao, Chen & Hu, 2000).

The population diversity of GA would be greatly reduced after some generations, and might lead to a premature convergence to a local optimum. Actually, it tends to converge prematurely and the optimization may get stuck at a local optimum. For example, the population is not always sufficiently huge in size to typical GA problem solving. In order to overcome these flaws, the key point is to maintain the population diversity and prevent the incest leading to misleading local optima (Syswerda, 1989; Eshelmen and Schaffer, 1991).

To maintain the population diversity of GA, the concept of chaos is introduced in this paper. Chaos being radically different from statistical randomness, especially the inherent ability to search the space of interest efficiently, could improve the performance of optimization procedure. Chaos can be considered as an irregular motion, seemingly unpredictable random behavior under deterministic conditions. Random and chaotic motions should be distinguished here by their features. The former is reserved for problems in which to know the input forces are not necessary, but some statistical measures of the parameters are enough. However, chaotic is reserved for deterministic problems in which there are no random or unpredictable inputs or parameters.

In chaos, a small difference in the initial conditions may produce an enormous error in the final phenomena. It is extremely sensitive to the initial conditions, and its property sometimes referred to as the instability in the so-called butterfly effect or Liapunove's sense (Lorenze, 1963; Kim and Stringer, 1992). Sensitive dependence on initial conditions is often exhibited by multiple elements with nonlinear interactions in the systems. Owing to chaos characteristic, the system could be designed as an efficient approach for maintaining the population diversity in the problem of interest.

Clustering is one of the most important and the most challenging of classifying algorithms. A successful clustering algorithm is able to reliably find true natural groupings in the data set. K-means is one of the well-known algorithms for clustering, originally known as Forgy's method (Forgy, 1965). K-means clustering is the process of dispatching a set of objects into groups or clusters of similarities. Objects collected in the same cluster have similar features, but others are not (Han & Kamber, 2001). K-means is famous for its simplicity and computational efficiency in clustering techniques. As aforementioned Chaotic algorithm is for population diversity in GA, and K-means is for convergence efficiency in evolution. The former will keep the system accuracy, and the later will decrease iteration times of GA significantly.

The remainder of this paper is organized as follows. In Section 2, gives an overview of the theorem and algorithm which will be encountered in this study later. In Section 3, a K-means clustering algorithm for Chaos GA is presented. And Section 4, KCGA is employed to search the optimization solution of a construction management issue. Section 5 provides some concluding remarks.

2. Theorem and Algorithm

2.1The Chaotic Concept and Logistic Mapping

Chaos can be considered traveling particles within a limited range occurred in a deterministic nonlinear dynamic system. There is no definite regularity for such a traveling path. Such a movement is very similar to a random process, but extremely sensitive to the initial condition. Chaotic dynamic mappings have been defined as noninvertible mappings of the (0, 1) interval onto itself. Logistic mapping (May, 1976; Feigenbaum, 1978) is one of the most important Chaotic dynamic mappings which defines the simplest mapping for studying the period-doubling bifurcation (vide infra). In the well-known logistic equation (May, 1976):

$$X_{n+1} = f(\mu, X_n) = \mu X_n (1 - X_n) \qquad (1)$$

in which μ stands for a control parameter, X for a variable and n =0,1,2,3…. It is easy to find that equation (1) is a deterministic dynamic system. The variable X is also called as chaotic variable. The basic characteristic of chaos could be presented by Eq. (1), for a very small difference in the initial value of X will cause large difference in its long-term behavior.

The variation of control parameter μ of Eq. (1) will directly impact the behavior of X greatly. Usually, [0, 4] has been defined as domain area of control parameter μ. Different value in domain area of μ will determine whether X stabilizes at a constant size or behaves chaotically in an unpredictable pattern. The track of chaotic variable looks like in disorder. However, it can travel ergodically over the whole space of interest especially under the condition of μ = 4. Then, a tiny difference in initial value of the chaotic variable would result in considerable differences of the values of chaotic variable later. Generally, there are three primary characteristics of the variation of the chaotic variable, i.e. ergodicity, irregularity and pseudo-randomness (Bountis, 1995; Li & Jiang, 1998; Ohya, 1998).

Logistic equation as shown in equation (1) can be distinguished by four intervals in accordance with the value of μ. First, when the value of μ is smaller than 1.0, the chaotic variable X_{n+1} converges to a stable point 0.0. Then, if the value of μ is between 1.0 and 3.0, no mater what initial value for X_0 between 0.0 and 1.0 was taken, X_{n+1} would converge to a certain value between 0.0 and 0.63665. And, the bifurcation occurs from μ≧3.0. The system will enter the chaos domain, if μ reaches a critical point of 3.5699456…. Finally, when μ =4.0 the values of X_{n+1} would take any real numbers between 0.0 and 1.0 and no redundant value would present again while having turned up already. In this study, 'μ =4.0' was taken to have the advantages of diversity during evolution.

2.2 The concept of GA and CGA

Genetic algorithms (GA) are designed by randomized search and optimization techniques. The principles of evolution and natural genetics are built in functions to GA accompanied with a large amount of implicit parallel features. GA contains a fixed-size population of potential solutions over the search space. The idea population could be created by an objective or fitness function or based on the domain knowledge of GA. These potential solutions are named individuals or chromosomes. GA consists not only of binary strings-individuals, but other encodings are also possible. For instance, in the literature (Wright, 1991; Michalewicz, Janikow & Krawczyk, 1992), a real-coded GA was proposed and the individual vector was coded as the same as the solution vector. The evolution usually starts from a population of randomly generated individuals and continued by selection, crossover, mutation in each iteration.

In every iteration, a new population is created and based on the following four steps:

(1) Evaluation: each individual of the population will be evaluated and assigned a value derived from fitness function.

(2) Selection: individuals with higher fitness value will be more likely to be selected for next generation. Here, a competitive strategy was used to selection to improve its performance.

(3) Crossover: the crossover process was to choose two individuals as parents randomly. This study applies one-point crossover process in which the point is randomly selected in the list of fields. All the fields lying after this point was exchanged between the two parents to create two new offspring.

(4) Mutation: The mutation process is a probability-based procedure in which a heuristic operation was employed to find shortest path from a random point. Then, a correction action is taken to keep individuals meeting the legal requirements, in case of necessary.

The above four steps are iterated in this study until a satisfactory solution is found or the terminating criterion is met.

In this study, while a crossover has finished, the new generated offspring may not follow the designed rule to visit every node once and move back to the starting point. A new offspring will compare with the swapped and original portion to verify if the members are identical. Same members lead to a sound crossover while duplicated members with parents need to be legalized. For instance, a one-point crossover was introduced; the random selected point of field is 3 and 2 shown on the following two tables.

Table 1 Legalization to crossover with identical members

Parents	Selected field	Swapping	Operation	Offspring
1 2 3 [5 4 1]	Crossover on field 3	1 2 3 [4 5 1]	Equal to	1 2 3 4 5 1
1 3 2 [4 5 1]		1 3 2 [5 4 1]		1 3 2 5 4 1

Table 2 Legalization to crossover with un-identical members

Parents	Selected field	Swapping	Operation	Offspring
1 2 [3 5 4 1]	Crossover on field 2	1 2 [2 4 5 1]	Legalization	1 3 2 4 5 1
1 3 [2 4 5 1]		1 3 [3 5 4 1]		1 2 3 5 4 1

To improve the performance of GA search, it should keep individuals scattered in the whole searching space. After adopting the nature of the chaotic process, a new GA search method was formed. Chaos-Genetic Algorithms (CGA), integrating GA with chaotic variable, was proposed in this work and would be improved by incorporating clustering techniques later. CGA holds both advantages of GA and the chaotic variable. It can keep the individuals distributed ergodically in the defined space and avoid from the premature of generations. And, CGA also takes the inherent advantage of GA over convergence to overcome the randomness of the chaotic process and hence to increase the probability of finding the global optimal solution.

2.3 The K-means Clustering concept in GA

Clustering is the process of grouping a set of physical or abstract items into clusters by similar features. K-means is one of the well-known algorithms for clustering, and it has been employed extensively in various fields including exploring studies: such as data mining, statistical data analysis: such as Custom Relationship Management, and other business applications. The K-means algorithm for clustering is based on the mean value of items in the group. It is suggested to assign each item to the cluster with the nearest centroid (mean) (Mac- Queen, 1967). In general, in this study the primary operating procedures for K-means are presented as follows:

(1) Defining how many clusters are to be created.
(2) Randomly assigning initial items to different clusters.
(3) Assigning new items to the cluster whose location to centroid is the nearest (by Euclidean distance with either standardized or un-standardized observations) and re-calculate the centroid for the existing or updated clusters.
(4) Repeating Step (3) until no more reassigning.

3. Proposed K-means clustering and Chaos in Genetic Algorithm

Assume that the working individual of independent variables is denoted by x consisting of n elements. They are named and denoted by x1, x2, . . ., xn . Thus, a problem of searching minimum could be described as:

Min f (x1; x2; . . . xn)
s.t. xi∈(ai, bi) i=1, 2,3, . . . , n (2)

Function f is related to the value of dependent variables x, which is subject to be optimized. The lower and upper limit of xi in function f are [a1, a2, ..., an] and [b1, b2, …, bn], respectively. The chaotic process could be defined through the following equation as the same as Eq. (1) (Li & Jiang, 1998; May, 1976):

$$cx_i^{k+1} = 4\,cx_i^{(k)}\,(1- cx_i^{(k)}) \quad i = 1, 2, ..., n, \qquad (3)$$

in which cx_i is the ith chaotic variable, and (k) and (k+1) denote the number of iterations. Then a linear mapping function was used to convert chaotic variable to a certain interval. In this study the linear mapping function could be described as:

$$x_i^{k+1} = a_i + cx_i^{(k)}(b_i - a_i) \quad i = 1, 2, ..., n, \qquad (4)$$

in which x_i^{k+1} is the *i*th working variable, and (k) and (k+1) denote the number of iterations. a_i and b_i are the lower and upper limits.

K-means plays a critical role in convergence of GA. Chaos algorithm can keep GA population diversity and avoid from premature. To take advantages of the above described benefits in GA, a novel algorithm combined K-means clustering and the Chaos algorithms with genetic algorithms was proposed as a powerful hybrid algorithm called KCGA (K-means and Chaos in Genetic Algorithm). Initial population of KCGA should be generated from chaotic algorithm, and then chaos function would adjust the individuals after mutation with the same proability. After mutation, K-means clustering in this study will help to group population in several clusters as pre-defined. Thus, location information of each centroid of cluster would be treated as candidate individuals for next generation. A competing procedure was employed to eliminate lower fitness value individuals, and reserved the others to create formal population for KCGA iteration.

During the convergence, GA generates a certain rule to direct population's migration. In particular, K-means was used with GA to thoroughly explore the entire search space so that to find out the most possible migration way and potential individuals for conventional GA. First, each individual in a population of GA denotes a set of feasible solution generated by chaos algorithm. Second, given all individuals as input, the K-means clustering algorithm can locate the centroid of each cluster. Third, the new formed centroids of each cluster will convert to candidate individuals appending to the existing population. These new formed centroids also indicate the moving centers of current iteration. Fourth, fitness values of individuals are evaluated and by a competing algorithm to keep enough individuals for next iteration. And, the flow chart of K-means Chaos Genetic Algorithm is described as following in figure 1.

4. Experimental result

Construction work includes many inherently hazardous conditions and tasks such as work at noise, dust, height, excavations, etc. For example, construction has about 6% of U.S. workers, but 20% of the fatalities - the largest number of fatalities reported for any of the industry sectors. These were announced by National institute for occupation safety and health (NIOSH) in 2008. In this study, a simulated case of ten building-construction sites was used for an auditor of safety and health. The auditor should start from one of the building-construction sites and travel to every site before returning back to the same place. The target is to find out the shortest path along every construction site.

After assigning each construction site an integer number, the distances between each site could be recorded and create a matrix. The fitness function was designed to calculate the total distance along the path. Any set of random integer number within [1, 10] may stands for a different path. To comply with the real world, it is critical to legalize offspring during the KCGA iteration, especially after the crossover and mutation procedure. Chaotic algorithm will impact the visit priority of each site. However, each centroid of cluster derived from K-means may not be integer. They need to re-legalize again to get their integer sequence as one of the quasi individuals for next generation. All experiments are completed on Core 2 CPU T5500 @ 1.66GHz PCs with 2GB memory. The results reported are all averaged over 50 independent runs. The parameters, such as mutation rate, crossover rate, generation limit, are given under the results.

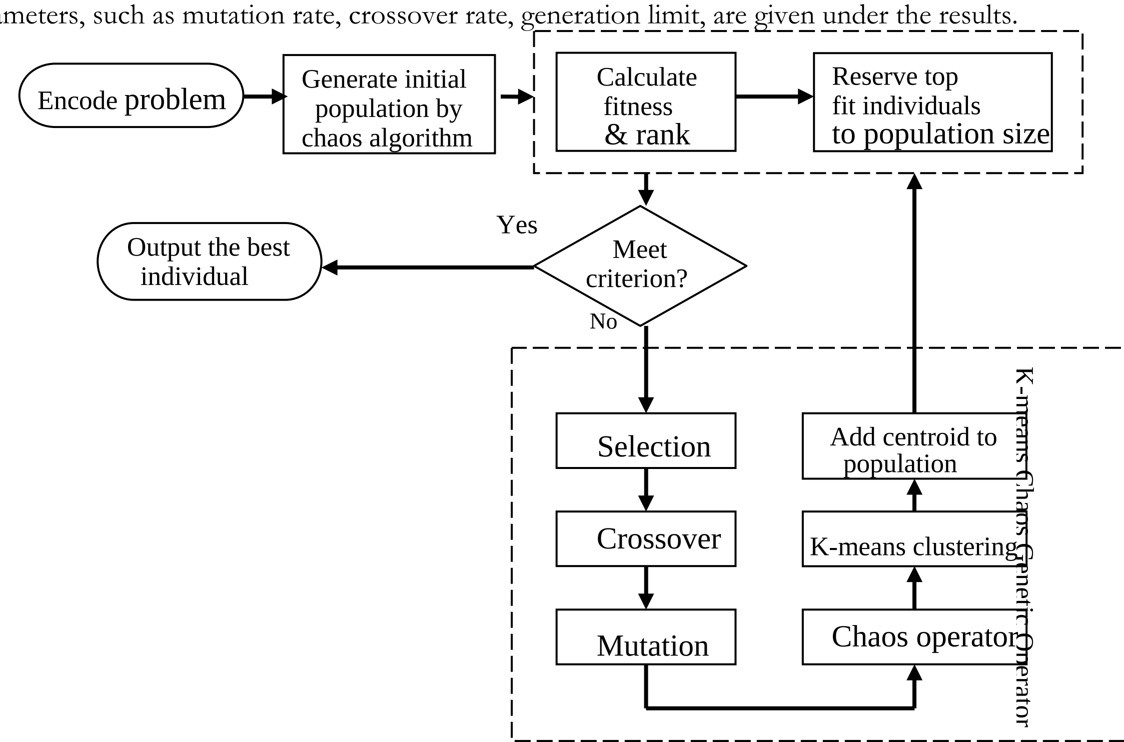


Fig. 1. Flow chart of K-means chaos genetic algorithm

From GA to KGA, a K-means clustering technique adopted by genetic algorithm can speed up its convergence rate by saving seventy percent of iterations and keep the accuracy over eighty percent. It is easy to find that the standard error of iteration has tremendously descended from GA (or CGA) to KGA in table 3 listed as above. The result has strongly recommended that a speeding convergence of searching in ten-dimension space can be smoothly realized by K-means clustering technique efficiently.

Combining K-means clustering technique with GA could assure to converge. It was shown that KCGA and KGA had never failed to converge during their experimental procedures for they had identical values in minimum and maximum, defined by GA as a criterion of termination.

GA, integrated with K-means clustering technique and chaos algorithm, could promote its accuracy and reduce the converging time. Migration from GA to KCGA, listed in table 3, has shown that KCGA improves the accuracy of GA, and diminishes the amount of evolution runs significantly.

Table 3 The performance of four methods

	KCGA (0.90)		KGA (0.80)		CGA (0.85)		GA (0.85)	
	Avg.	Std. E.	Avg.	Std. E.	Avg.	Std. E.	Avg.	Std. E.
Iteration	23.9	5.0	31.0	38.3	71.8	132.5	105.0	171.0
Time(sec.)	1.3	0.3	1.5	2.0	1.1	1.6	1.1	1.5
Min.	38.1	0.3	38.3	0.6	38.1	0.3	38.1	0.2
Max.	38.1	0.3	38.3	0.6	38.2	0.5	39.2	3.0
Fitness	2286.0	18.5	2295.0	33.0	2289.0	22.0	2291.0	26.8

Notes: mutation rate = 0.01, crossover rate = 0.8, population size = 60,

generation limit = 500, Avg.: Average, Std. E.: Standard Error

(*) = accurate ratio

5. Conclusions

This study has proposed a procedure which joins K-means and chaos attributes based on genetic algorithm. The proposed procedure is not only to enhance the diversity of GA for more accuracy but also to extract clustering rules for achieving a potential trend of evolution. Additionally, it can effectively improve some drawbacks of traditional GA, such as long running time and getting trapped in local optima. Furthermore, this proposed procedure can really contribute to construction management in real world.

References

[1] Bountis, T. (1995). Fundamental concepts of classical chaos. Part I. Information Dynamics and Open Systems, 3 (23) 23_/96.

[2] Feigenbaum, M.J., (1978). J. Stat. Phys. 19, 25.

[3] Forgy, E. (1965). Cluster analysis of multivariate data: Efficiency versus interpreability of classifications. Biomertrics, 21, 768.

[4] Han, J., & Kamber, M. (2001). Data mining: Concepts and techniques. San Francisco: Morgan Kaufmann Publishers.

[5] Hibbert, D.B., (1993). Chem. Intel. Lab. Syst. 19, 277.

[6] Holland, J.H., (1975). Adaptation in Natural and Artificial Systems. University of Michigan, Ann Arbor.

[7] Hugget, A., Se´bastian, P., & Nadeau, J. P. (1999). Global optimization of a dryer by using neural networks and genetic algorithms. American Institute of Chemical Engineering Journal , 45 (6) 1227_/1238.

[8] Kim, J.H., Stringer, J., (1992). Applied Chaos. Wiley, New York.

[9] Li, B., & Jiang, W. S. (1998). Optimizing complex functions by chaos search. Cybernetics and Systems: An International Journal 29, 409-419.

[10] Lorenze, E.N., (1963). J. Atmos. Sci. 20, 130.

[11] MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate observations. In Proceedings of 5th berkeley symposium on mathematical statistics and probability (pp. 281–297). Berkeley: University of California Press.

[12] May, R., (1976). Simple mathematical model with very complicated dynamics. Nature 261, 45–67.

[13] Michalewicz, Z., Janikow, C. Z., & Krawczyk, J. B. (1992). A modified genetic algorithm for optimal control problems. Computer and Mathematical Applications 23 (12), 83_/94.

[14] Ohya, M. (1998). Complexities and their applications to characterization of chaos. International Journal of Theoretical Physics , 37 (1) 495-505.

[15] Syswerda, G., (1989). Uniform in Genetic Algorithms, ICGA' 89. Kaufmann (Morgan), California, p. 2.

[16] Wright, A. H. (1991). In G. J. E. Rawlins (Ed.), Genetic algorithms for real parameter optimization. Foundations of genetic algorithms (pp. 205_/218). California: Morgan Kaufmann.
[17] Zhao, W. X., Chen, D. Z., & Hu, S. X. (2000). Optimizing operating conditions based on ANN and modified gas. Computers and Chemical Engineering 24, 61/65.

An Object-oriented Framework for Spatial Interpolation

Yo-Ming Hsieh[1], and Mao-Sen Pan[2]

[1]Department of Construction Engineering, National Taiwan University of Science and Technology, No. 43, Sec. 4, Keelung Rd., Taipei, Taiwan; PH(886) 2-2730-1056; FAX(886) 2-2737-6606;
e-mail: ymhsieh@mail.ntust.edu.tw
[2]Department of Construction Engineering, National Taiwan University of Science and Technology, No. 43, Sec. 4, Keelung Rd., Taipei, Taiwan; e-mail: D9505503@mail.ntust.edu.tw

Abstract

Interpolation is an important operator in numerical methods for solving partial differential equations and in geospatial applications. There are many interpolation methods proposed in the past. In this work, a unified software framework is proposed through the use of design-patterns in object-oriented programming. By using this framework, little effort is necessary to implement different interpolations algorithms when commonality with implemented algorithms can be found. Furthermore, through this framework, it becomes easy to compare the performance of different algorithms because of the unified application interface..

1. Introduction

Spatial interpolation has wide applications in the field civil engineering, e.g. solving PDE (partial differential equations) and GIS (geographical information systems). Currently, there is neither standard nor widely adopted API (application programming interface) for interpolation. As a result, it is necessary for developers of aforementioned applications to implement their own interpolation operators or adapt their applications to some developed codes. If ten interpolation methods is to be evaluated for understanding its applicability or its performance, the developer suffers because they need to read 10 different documents, they may have different data interfaces, and they may even conflict in names. Therefore, there is a dare need for a standard interface for performing interpolation.

A software framework for interpolation is proposed in this work to unify implementations of different interpolation algorithms. The unification is made possible by the use of encapsulation, inheritance and *polymorphism* characteristics of OOP (object-oriented programming). The design is guided by design patterns, a concept pioneered by Gamma(1991) to achieve low coupling between different classes.

It is believed the proposed software framework can potentially benefit developers of both 1) new interpolation algorithms by encouraging code reuse; and 2) applications in need of interpolation by having a unified interface to multiple interpolation algorithms.

2. Methodology

This study follows the flowchart shown in Figure 1. It is seen in the flowchart that this study consists the following steps: survey, common feature extraction, core class design, system interface design, and evaluation.

First step of this study was a survey of interpolation methods interested. In this survey, few methods for interpolation commonly used in solving PDE were included. These methods were selected solely because these were needed by other studies in author's research group, but authors believed the generality of the proposed software framework should not be affected much by the limited scope of the survey.

Common procedures and requirements were then analyzed and extracted from the surveyed interpolations algorithms. This step is necessary in order to identify necessary common function call interfaces and common data interfaces. This step contributes to the design of a general software framework that unifies the implementation of interpolation algorithms.

From the identified common requirements between interpolation algorithms, abstract core classes were then designed. The abstract core classes define the common interface, including both data interface and API interface, for implementing interpolation algorithms. The data interface was carefully designed with great generality to ensure it satisfies all the needs of surveyed algorithms; the API interface or class collaboration were designed with aids from design patterns to lower the coupling between classes in order to maximize extensibility while minimizing changes needed when adding new algorithms.

In order to make the framework for interpolation easy to use for application developers, a wrapper class or system interface class is then designed to wrap up the above mentioned core classes to provide a simple to use application interface for application developers. More details about the system interface class are given at later paragraphs.

Once all classes were designed, the designed classes were validated and evaluated by implementing surveyed algorithms in order to make sure these classes: 1) can implement all the surveyed methods, and 2) encourage good code-reuse.

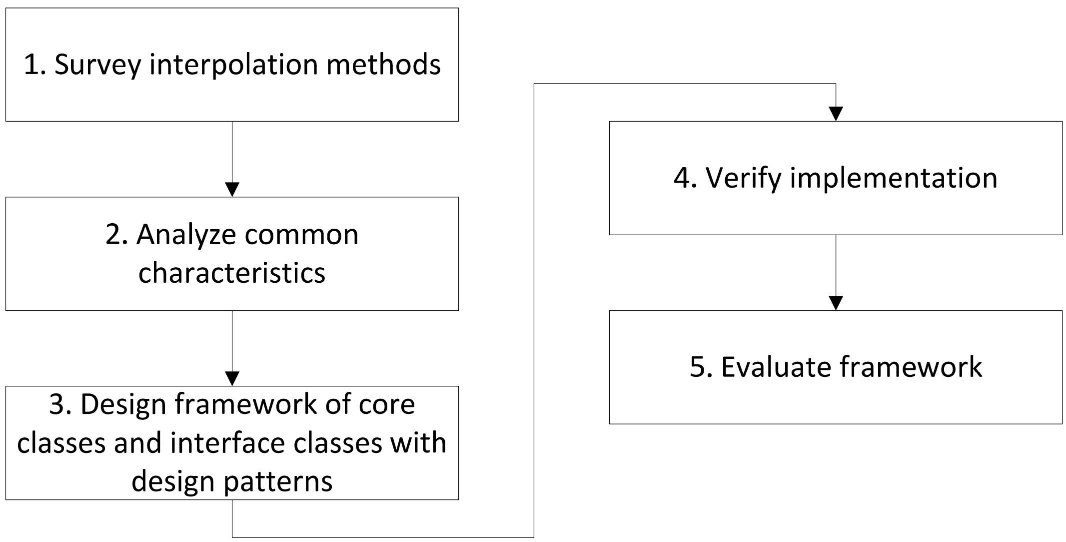


Figure 1: The overall flowchart of this study

3. Core Interpolation Framework

Before the proposed software framework is described, interpolation is first defined. In this work, spatial interpolation is defined as an evaluation of a quantity of interest at a given spatial location based on the quantities of the same type that exists on a set of known points.

Based on surveyed interpolation algorithms, it was determined the evaluation of one interpolation operation often involves evaluations of: 1) weighting, which controls the how influential each known point is to the point of interest; 2) basis function, which controls how the quantity to be evaluated is distributed continuous in space; and 3) neighboring points, which finds surrounding points for a given point.

Accordingly, four abstract classes were designed to define interfaces for the aforementioned four evaluations. These classes need to be derived or "sub-classed" to define concrete implementation of these evaluations. The collaboration of these classes is shown in Figure 2, and is discussed in the following paragraphs.

The *interpolation* class is the main "driving" class for performing interpolation. Applications in need of interpolation will create a concrete object derived from this type. This interface class defines the common interface for initializing data required to perform interpolation and performing interpolation for all interpolation algorithms. Figure 2 shows an example of implementing *MLS* (moving least square, Lancaster and Salkauskas 1981) interpolation algorithm. The *MLS* class is a concrete class inherited from *interpolation* class, and the concrete class overrides data initialization method and provides an implementation of MLS interpolation algorithm. In a way, the interpolation class is similar to a template class that defines the collaboration between *weightFunction*, *basisFunction*, and *searchAlgorithm* abstract classes. The class itself does not assume or use any concrete class. Therefore, this design ensures great extensibility or flexible for implementing interpolation algorithms. Furthermore, if a particular interpolation method does not use *weightFunction*, *basisFunction*, or *searchAlgorithm* classes, the implementer can simply ignore references to the aforementioned three abstract classes and use only the defined interfaces by *interpolation* class.

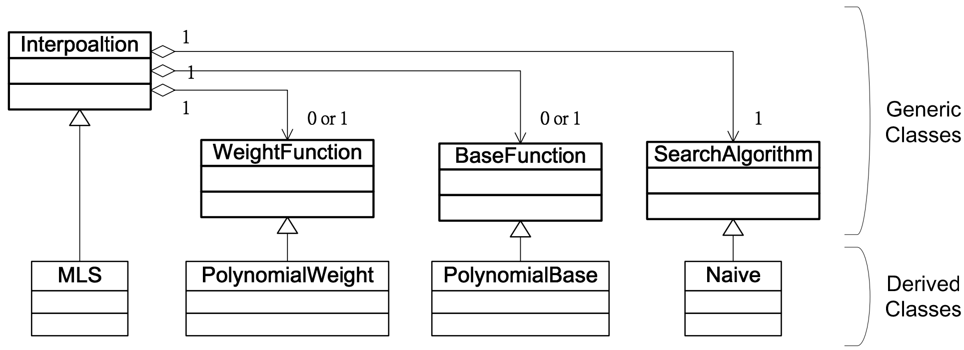


Figure 2: Class collaboration for performing interpolation

The *weightFunction* class, similar to the *interpolation* class, is an abstract class that defines universal interface for determining weights of existing points with known quantities to the point to be interpolated. There were many weighting functions proposed in the past, e.g. Liu(2003), Li et al(2004). If a weighting function is to be implemented based on the proposed framework, they need to sub-class the abstract *weightFunction* class to ensure that they can be easily incorporated or "bridged" into any interpolation algorithm.

The *basisFunction* class defines the common interface for basis functions (also known as interpolation function, shape function, kernel function, etc.), such as polynomials and radial basis functions (Liu 2003; Liu and Gu 1999; 2001; 2005; Wang and Liu 2000). Any implementation of basis function evaluation needs to sub-class from *basisFunction* class and implements the virtual functions for calculating local interpolation.

Finally, the *searchAlgorithm* class is a helper abstract class for searching a specified number of neighboring points from a given set of points. It is considered as a helper class because interpolation algorithms usually do not concern how to find neighboring points, but most interpolation algorithms do need to find surrounding points for a given spatial location. It should be noted the efficiency of search algorithm tends to dictate the interpolation efficiency, and efficient search algorithms is necessary in order to get performing interpolations.

Two design patterns, bridge and strategy, were used in the design of class collaborations shown in Figure 2. The bridge design pattern "*decouples abstraction from implementation so that implementation and abstraction can be varied independently*" (Gamma et al. 1995). Derived classes of *interpolation* are to be implemented by using abstract interfaces of other three classes that are aggregated to the *interpolation* class. Therefore, any concrete implementation of *interpolation* can use any combination of concrete classes derived from *weightFunction*, *basisFunction*, and *searchAlgorithm*. Furthermore, the *interpolation* class itself is an abstract class, which can be refined further to develop yet another abstract class if desired.

The other design pattern applied in the design is the strategy design pattern. The strategy design pattern is a behavior design pattern that allows algorithms to be swapped at runtime. This is achieved by defining a virtual function that needs to be implemented by all sub-classes of abstract core classes such as *interpolation*, *weightFunction*, etc. Therefore, applications are allowed to choose which concrete sub-classes of *weightFunction*, *basisFunction*, and *searchAlgorithm* are to compose the *interpolation* class.

The core interpolation framework benefits developers of interpolation algorithms. With loose-coupling between concrete classes, it is easy to reuse existing classes and vice-versa. Efforts of implementing new interpolation algorithms that shares common weight functions or basis functions can be reduced. By using the framework, many different interpolation algorithms can be implemented by different composition of weighting, basis, and interpolation classes. On the other hand, the implementation of search algorithm dictates the efficiency of interpolation and should not be overlooked.

4. Interface Class

The core interpolation framework benefits developers for interpolation algorithms by encouraging code-reuse with well organized program structure. However, to ensure the maximum flexibility of the framework, it was decided that instantiation of concrete classes in the core framework is the responsibility of the program which uses the framework. This may be inconvenient for application developers who solely want to perform interpolation without knowing too much detail regarding interpolations. Therefore, an interface class following the "Façade" design pattern is designed to ease the use of core interpolation framework.

Figure 3 shows the composition of this interface class with the core interpolation framework. The interface class is responsible for instantiating concrete classes of *interpolation*, *weightFunction*, *basisFunction*, and *searchAlgorithm*, and then initializes each of these classes properly. One of the most important responsibilities of this class is to make sure the "right" concrete classes of *weightFunction*, *basisFunction*, and *searchAlgorithm* are assigned to the concrete *interpolation* class.

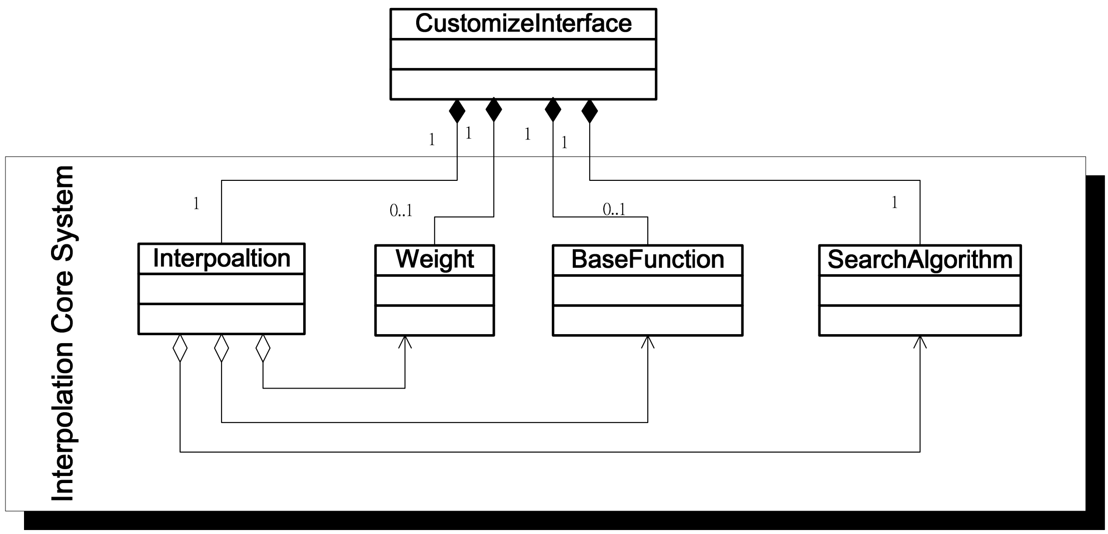


Figure 3: Interface class for the core interpolation framework

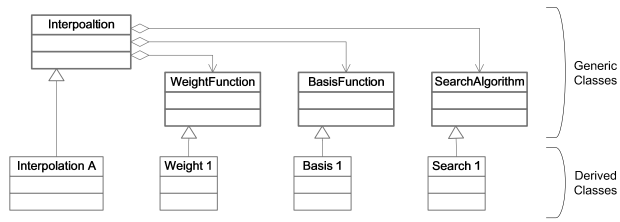
By introducing the interface class, application programs only need to know and use the interface class without knowing it is supported by four classes (*interpolation*, *weightFunction*, *basisFunction*, and *searchAlgorithm*). Therefore, the interface class benefits application developers who need simple interpolation operations without knowing details of interpolation algorithms.

5. Evaluation

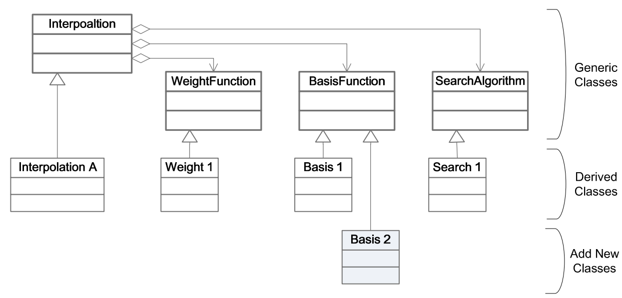
The designed software framework is evaluated from two perspectives: interpolation algorithm developers and application developers. For interpolation algorithm developers, the framework should reduce the effort of programming; for application developers, the framework ought to provide easy-to-use interfaces.

Figure 4 shows the benefit of the software framework for interpolation algorithm developers. Assuming an implementation for interpolation method A has been completed, as in Figure 4(a). One may create another interpolation method B, as in Figure 4(b), by 1) implementing a new concrete *basisFunction* class and 2) creating another concrete *interface* class that aggregates new *basisFunction* class with old concrete *weight interpolation*, and *searchAlgorithm* classes. Similarly, different interpolations can be realized by substituting weight functions, as in Figure 4(c). Other interpolation methods are possible by using the same weight and basis functions but different interpolation procedures, Figure 4(d). Finally, if one is unsatisfied with the performance of interpolation, he or she may try to improve the performance by introducing better search algorithms, as in Figure 4(e). This great flexibility or extensibility is attained by *encapsulation* and *inheritance*, two important characteristics of OOP.

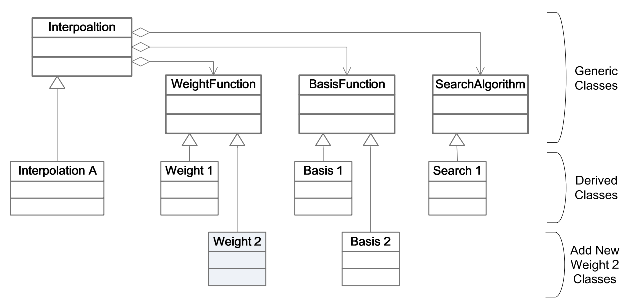
Figure 5(a) shows the ease of maintaining interface class. User can also define the interface class himself. Basically, once a new interpolation algorithm is implemented, a new interface class is added by 1) programming constructor to instantiate four concrete classes of appropriate combination; and 2) writing an inline method that calls to the calculation method of the concrete interpolation class. It may be noted that identical combination of core interpolation classes with different "default" parameters may be programmed into a different interface class with different parameter list under the same method name – an example of using *polymorphism*.



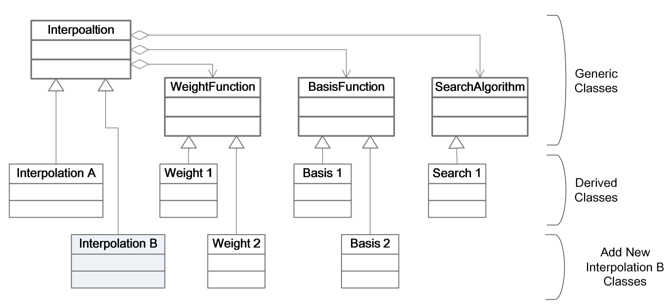
(a) Assuming interpolation method A has been completed



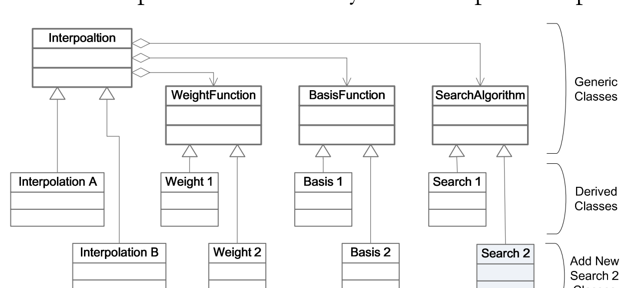
(b) New interpolation algorithm with new basis function



(c) New interpolation method with new weight function



(d) New interpolation method by new interpolation procedures



(e) Gaining better interpolation efficiency by implementing efficient search

Figure 4: Benefits for interpolation algorithm developers

Figure 5(b) shows the effort required for application developers who need performing interpolation, assuming core interpolation classes have been implemented and bundled as a class library. It is seen that the effort involves 1) preparing data, 2) instantiating a concrete interface class, and 3) calling the calculation method of the interface class. These efforts are necessary regardless the practice of programming. Therefore, it is considered by authors that the design of the interface class is essential to help application developers who do not concern about details of interpolation algorithms.

Figure 5(c) shows the ease of switching to different interpolation algorithms using the interface class. This is sometimes necessary because application developers do not necessarily know beforehand which interpolation algorithm suits his application the most. By changing one line of code, which calls to the constructor of a particular constructor of a concrete interface class, the interpolation algorithm can be completed altered – achieved through *inheritance*.

```
//====================//
//  Interpolation interface header  //
//====================//

//In this area include necessary header file. (interpolation method, weight function …)
#include <xxxx.h>
…

enum WeightType {_noWeight, _weight1=1, _weight2, … };
enum BasisType {_noBasis, _basis1=1, _basis2, …};
enum SearchType {_noSearch, _search1=1, _search2, …};
enum InterpolationType {_noInterpoaltion, _interpoaltion1=1,
        //Enumeration all types methods for weight, basis, search, and interpolation methods. It
        needs to update enum parameters if new algorithms be added.

class InterpolationInterface {
protected:
    //basic field  parameters
    int dimension;
    …

    WeightFunction    *weightFunction;
    BasisFunction       *basisFunction;
    SearchAlgorithm  *search;
    Interpolation        *interpolation;
        //Programming maintainer use these pointer to get default or selection algorithms objects.

    void createInterpolation(InterpolationType I
    void createWeight(WeightType weight);
    void createBaseFunction(BasisType base);
    void createSearch(SearchType search);
        //Programming maintainer use these functions to realize objects. These Functions get parameter which selected
        from enumeration types, then create objects. It also needs to update if new algorithms be added.

    void InterpolationBuilder();
        //Combining appropriate algorithms in this function.
public:
    // step 1. Constructor and Destructor
    InterpolationInterface(basis parameter1, basis parameter2, …);
    ~InterpolationInterface ();

    // step 2.User select one of below method to build interpolation ob
    void create(parameter1, parameter2, …);  //create by advance use
    void createInterpolationAlgorithm1();       //default method
    …

    // 3. User use this function to calculate and get interpolation resul
    void calcField(calc parameter1, calc parameter2, …);
        //Programming maintainer provide interface for user create interpolation. User can use default method function
        to create easily or give selection parameters by advance function to create.

};
```

(a) Maintenance of interface class

Figure 5: Benefits of proposed framework for application developers

```
// step 1: creating use of interface class
      InterpolationInterface  *myInterpolation = new InterpolationInterface(parameter1, …);

// step 2: choosing one of style method for building interpolation object.
      myInterpolation -> createInterpolationAlgorithm1();

// step 3: calling the calculation method of use interface class.
      myInterpolation -> calcField(calc parameter1, calc parameter2, …);
```

(b) Use of interface class in applications

```
// step 1: creating use of interface class
      InterpolationInterface  *myInterpolation = new InterpolationInterface(parameter1, …);

// step 2: choosing one of style method for building interpolation object.           //Switching to different
      myInterpolation -> create(InterpolationAlgorithm1, basis1, weight1, searc    interpolation algorithms jus
                                                                                   change creation function of
// step 3: calling the calculation method of use interface class.                   step2.
      myInterpolation -> calcField(calc parameter1, calc parameter2, …);
```

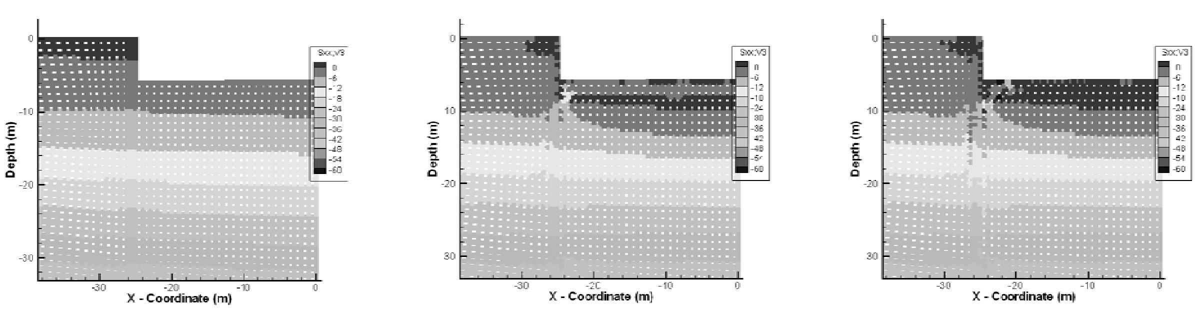
(c) Changing interpolation algorithms in applications

Figure 5 (cont'd): Benefits of proposed framework for application developers

6. Applications

The interpolation library developed based on the described framework was applied to evaluate various interpolation algorithms for stress interpolation in excavation analyses. These interpolations are necessary for various purposes such as solving partial differential equations using mesh-free methods and post-processing analysis results for producing contours, etc.

Figure 6 shows interpolations of horizontal stresses produced by an excavation analyses. It is seen that different algorithms may produce different results due to differences in algorithms, proximities to boundaries, etc. Therefore, it is necessary to have the proposed interpolation framework to assist evaluations of various interpolation algorithms to find the most-suitable interpolation algorithms for one's own purpose.



(a) Least-Square (b) Moving Least-Square (c) PIM

Figure 6. Different stress-distributions by different interpolation algorithms

7. Summary

A software framework for implementing interpolation algorithms is proposed based on object-oriented programming and well-known design patterns. The framework is considered general and should be capable of unifying all implementations of interpolation algorithms. The proposed framework benefits implementers of interpolations by providing great extensibility, good code-reuse and code-management; it also benefits application developers by the ease of use and ease of switching to different algorithms.

Acknowledgement

This work is partially supported by NSC 95-2221-E-011-110 of National Science Council of Taiwan.

References

[1] Gamma, E. (1991). "Object-oriented software development based on ET++: design patterns, Class Library, Tools." PhD thesis, University of Zurich Institut für Informatik.
[2] Gamma, E., Helm, R., Johnson, R., and Vlissides, J. (1995). "Design Patterns Elements of Reusable Object-iented Software." Addison-Wesley, USA, ISBN 0-201-63361-2.
[3] Lancaster, P. and Salkauskas, K. (1981). "Surfaces Generated by Moving Least Squares Methods." Math. of Comp., 37(155), 141-158.
[4] Li, H., Wang, Q. X. and Lam, K. Y. (2004). "Development of a novel meshless Local Kriging (LoKriging) method for structural dynamic analysis." Comp. methods in Appl. Mech. And Engrg., 193, 2599-2619.
[5] Liu, G. R. and Gu, Y. T. (1999). "A point interpolation method." Proceedings of 4th Asia-Pacific Conf. on Comp. Mech., Singapore, 1009-1014.
[6] Liu, G. R. and Gu, Y. T. (2001). "A point interpolation method for two-dimensional solids." Int. J. Muner. Meth. Engng, 50, 937-951.
[7] Liu, G. R. (2003). "Mesh Free Methods – Moving beyond the Finite Element Method." CRC Press, ISBN 0-8493-1238-8.
[8] Liu, G. R. and Gu, Y. T. (2005). "An introduction to meshfree methods and their programming." Springer, ISBN 1-4020-3228-5.
[9] Wang, J. G. and Liu, G. R. (2000). "Radial point interpolation method for lastoplastic problems." Proc. of the 1st Int. Conf. On Structural Stability and Dynamics, Taipei, Taiwan, 703-708.

Sparse Reconstruction and Geo-Registration of Site Photographs for As-Built Construction Representation and Automatic Progress Data Collection

Mani Golparvar-Fard[1], Feniosky Peña-Mora[2] and Silvio Savarese[3]

[1] William E. O'Neil Pre-Doctoral Candidate in Construction Management and Master of Computer Science Student, University of Illinois, Urbana-Champaign, 205 N. Mathews Ave., Urbana, IL 61801; PH (217) 333-2071; FAX (217) 265-8039; email: mgolpar2@illinois.edu

[2] E.W. and J.M. Gutgsell Endowed Professor in Construction Management and Information Technology, Civil and Environmental Engineering, University of Illinois, Urbana-Champaign, 205 N. Mathews Ave., Urbana, IL 61801; PH (217) 244-0187; FAX (217) 265-8039; email: feniosky@illinois.edu

[3] Assistant Professor, Electrical and Computer Engineering, University of Michigan, Ann Arbor, MI 48109-2122; PH (734) 647-8136; FAX (217) 265-8039; e-mail: silvio@eecs.umich.edu

Abstract

Most of the current techniques for automating progress data collection promise to eliminate labor-intensive tasks associated with manual data collection. A drawback is the necessity to add additional steps to be performed before, during, or after utilization of such technologies. Working with such featureless data and without having semantic information of the scene, geometric-reasoning is problematic and induces estimation errors. In this paper application of unordered daily progress photograph logs, available on any job site, as a data collection technique is explored. In our proposed approach, a sparse 3D geometric scene of a construction site is reconstructed and photographs are geo-registered. This allows project managers to remotely explore as-built scene and geo-registered site photographs at different stages of progress, minimize their travel time, perform remote as-built analysis and use the proposed system as a tool for contractor coordination purposes. Furthermore, the point cloud allows the planned model to be registered with the as-built scene, in turn supporting development of the automatic 3D recognition technique and quantification of as-built progression from the geo-registered images. We present our results on two ongoing construction projects and further discuss technical issues on developing and implementing this technology for automation and visualization of as-built construction.

Introduction

In today's economy, construction companies are seeking new ways to streamline their work processes to reduce project durations and costs. The reasons are simple: owners need to minimize time and cost for their services or marketing products and thus need to reduce the delivery time for facilities that provide such services or products. Along the same line, contractors are faced with intense competition, tight market constraints and slim profit margins. These situations motivate contractors to detect actual or potential delays and cost overruns in field activities as early as possible. Systematic and comprehensive tracking and monitoring of construction performance, workforce productivity, site layout and quality provides managers with an opportunity to detect such delays and overruns, initiate remedial actions and increase the chance of controlling their impacts.

Despite the importance, current practice of data collection experiences several process inefficiencies. Every day, superintendents and field engineers must collect and extract extensive amount of as-built data (Navon and Sacks 2007) which in turn may cause human-errors and induce error in the manually collected data (observed by authors). The excessive load of the required work usually makes monitoring task non-systematic and leaves it to be based largely upon judgments derived from past experiences. This may create a tendency to let project plan inputs be used as performance measures which in turn affects quality of the results (Meredith and Mantel 2003). Some of the currently used techniques such as cost-based monitoring may create a time-lag between the time that actual progress is reported and the time that progress is actually obtained. Also site activities are more numerous than what plans usually describe. Consistent changes on the job site including location of construction equipment, workforce, materials and the work sequences usually not included in the original plans are left to be noted in text or chart forms (Shih et al. 2004). These recording and reporting forms may increase the time required to describe and explain the as-built situation, consistently changing layouts and constructability issues in coordination meetings as well as delaying the decision-making process (Golparvar-Fard et al. 2006). In summary, with current *data collection*, *analysis* and *reporting* methods it may not be easy to clearly and quickly understand the progress situation. Functions that enable automatic digital recording, identification and reporting of the as-built construction site will be useful.

Most of the current technologies for automating data collection (such as laser scanners, Radio Frequency Identification (RFID) tags, Global Positioning Systems (GPS), Wireless Fidelity (Wi-Fi) and Ultra Wide-Bands (UWB) sensors are promising if one wishes to eliminate labor-intensive and non-value adding tasks associated with manual site data collection. However, a drawback in application of these technologies is the necessity to add additional steps needed to be performed before, during, or after utilization of such technologies at a construction site (Kiziltas et al. 2008). For instance, by using laser scanners, only Cartesian coordinate information of the scanned scene could be retrieved. Working with such featureless data and without any semantic information of the scene, geometric reasoning based on this data is problematic and induces estimation errors.

In this paper, application of unordered daily progress photograph logs - which is currently available on almost any construction site - as a data collection technique is explored. Nowadays site photographs are becoming valuable sources of accurate project information (Soibelman et al. 2008). It is very common among all parties involved in projects (from construction managers to subcontractors and clients to architects) to take digital photographs from construction sites to create a complete progress photo-log and utilize the log for coordination and communication purposes as well as collecting them as supplementary documents for potential claims. Cameras can cover significant areas of a construction site, especially if outfitted with zoom lenses. They have the capability of providing positioning information about multiple construction entities concurrently. All of these facts indicate that project photographs have evolved into a significant and irreplaceable part of project documentation and thus providing solid participations for their usage as visual, real-time as well as easy to obtain, low-price data capturing technology which does not need any expertise. The availability of such rich imagery of large parts seen under different viewing conditions motivated this study to see how based on this valuable dataset, digital representation of the as-built site can be generated, allowing progress to be tracked and workspace logistics, constructability, quality, safety, as well as productivity to be analyzed.

In our proposed approach, a sparse 3D geometric scene of the site is reconstructed and progress photographs are geo-registered in a virtual environment. This allows project managers to interactively and remotely browse and explore the as-built scene and geo-registered site construction photographs in a 3D environment. We show from the stand point of progress monitoring how these site photograph logs present an ultimate data set, giving the ability to model a significant portion of as-built geometry at high resolution respective to conditions where enough photographs are being taken. Within the proposed platform, automatic 3D recognition techniques could be developed to quantify as-built progress from the geo-registered images. We present our results on two ongoing construction projects and further discuss technical issues on developing and implementing this new technology for generating and visualizing as-built scenes.

Emerging Field Data Capture Technologies

For more than a decade, researchers have been pointing out deficiencies in current construction site data collection practices (e.g., manual data collection, need for systematic collection and processing of as-built data to produce useful and real-time progress information (Kiziltas et al. 2008, Navon and Sacks 2007). According to (Navon and Sacks 2007) these research efforts have been motivated by two major drives: (a) an increasing need for real-time feedback and monitoring information and (b) rapid and cost effective technological development in automated data collection technologies for construction. The main technologies designed and implemented in this category are barcode and RFID tags, GPS Systems, Laser scanners and Time-Lapse Photography and Videotaping:

• Barcode and RFID tags have been used to capture and transmit data from a tag embedded or attached to construction components (Kiziltas et al. 2008, Navon and Sacks 2007). Unlike barcodes, RFID tags do not require line-of-sight, close proximity, individual reading and direct contact (Kiziltas et al. 2008). RFIDs and barcodes potentially eliminate non-value adding tasks associated with project management processes, but they require frequent installation and maintenance. Additionally they cannot be attached to many types of components and they do not capture progress of partially installed components.

• Laser scanners have been used for construction quality control (Akinci et al. 2006, Jaselkis et al. 2006), condition assessment (Gordon et al. 2004), component tracking (Teizer et al. 2005) and progress monitoring (El-Omari and Moselhi 2008, Bosche and Haas 2008, Su et al. 2006). Although laser scanners are promising in automating data collection, there still is a set of challenges in implementing such technology on construction sites. These limitations include discontinuity of the spatial information, mixed pixel phenomenon (Kiziltas et al. 2008) as well as scanning range and sensor calibration. For example, any moving object in line-of-sight of the scanner would not allow the point cloud of the under-study object to be captured. In addition, the moving object creates additional effort of the user to manually have the point cloud fixed. Also as the laser scanner gets away from the objects, the level of details within the captured components is reduced. Besides laser scanners require regular calibrations as well as warm up time. These limitations are parts of the time consuming process of data collection; however since laser scanners only provide Cartesian coordinate information of the scanned scene, processing such data is time consuming. Also they do not carry any semantic information, such as which point belongs to what structural components. Working with this type of featureless data makes geometric reasoning based on this data tedious and error prone (Kiziltas et al. 2008). Also none of these techniques provide any visual, reliable information about work sequence, site logistics or construction crew. Recently El-Omari and Moselhi (2008) presented a new interesting approach for data collection by combining 3D laser scanners and photogrammetry. The method was shown to be less time-consuming and has higher cost savings compared to single application of laser scanners. Their suggested approach minimizes access limitations of scanner placement, but the processing time required for each scan is still considerably high and the registration of images and 3D point cloud needs further adjustments. Also, laser scanners may not give the possibility of aligning site images taken from arbitrary view points with the 3D point cloud; yet in El-Omari and Moselhi (2008) the common points between laser scanner's 3D point clouds with images have been selected manually. Manual selection of common points between each image and point cloud may make such systems difficult to manage.

• GPS (Geographical Positioning Systems) as a location tracking tool also need line-of-sight between the receiver and the satellite; therefore it cannot normally operate indoors limiting the project context that could be monitored. Behzadan et al. (2008) suggests using WLAN technique as a tracking technique for indoor locations but they also report difficulties in using WLAN set ups on actual construction site and they relate these efficiencies to ongoing works (i.e., changes in soil, structure, plant and equipment, site layout). These inefficiencies necessitate WLAN system to be calibrated after regular intervals to maintain a high level of accuracy. Such regular calibration requirements make the system difficult to manage.

• Time-Lapse Photography and Videotaping: Previous research efforts in using time-lapsed photographs for the purpose of progress tracking goes back to Oglesby et al. (1989) wherein it was reported that application of site photographs allows analysts to focus on the details of the work face while being away from site tensions and confusions and perform time-studies on time-lapsed photographs for productivity improvement. However, lack of advanced technologies for automation had made the process time-consuming and unattractive to some extent. More recently, Abeid et al. (2003) presented Photo-Net II wherein time-lapse digital movies of construction activities were linked with critical path activities. In Photo-Net II, time-lapse photography was used as a source of spatial as-built information. In addition, Golparvar-Fard et al. (2007) also recently presented an Augmented Reality (AR) system wherein 3D models are superimposed over time-lapsed photographs.

• Other techniques such as wearable computers (PDAs), speech recognition and touch screens have also helped to collected construction site data electronically (Reinhardt et al. 2000), but current systems still need full time observer(s) to input and process information (Navon and Sacks 2007) and have not minimized the time required to process data.

Also none of these techniques besides (El-Omari and Moselhi 2008) - in which photographs are used to provide more information about the context of the scene - provide visual reliable information about work sequence logistics, site layout or construction crews. Our approach considers all the aspects of as-built data collection: *collection*, *analysis*, *communication* and *reporting*. We have looked into construction site photo-logs as existing simple yet robust data collection and communication techniques available on all construction sites and see how we can effectively use such information to address mentioned problems. In the section that follows, our proposed method of as-built data collection and representation is discussed.

Overview of the Proposed As-Built Representation

To date, application of unordered daily site photographs for representation of construction progress is almost unexploited. The progress images are usually not organized, uncalibrated, are widely variable and taken under various illumination, resolution, and image qualities. Developing computer vision and image processing techniques that can operate effectively with such imagery is a challenging task. Within such scope, one key challenge is image registration, i.e., gauging correspondences between images, and how they relate to one another in a common 3D coordinate system. This procedure is called Structure from Motion (SfM). While substantial progress has been done in these areas over the last two decades (recent examples: Snavely et al. 2008, Brown and Lowe 2005, Hartley and Zisserman 2004), some challenging aspects are still unresolved. For instance, there is a necessity to work with images that are capturing sites whose appearance is constantly changing due to progress or excessive movement of objects (e.g., site crew and machinery). Also site photographs are sometimes taken only from specific activities under progress or in panoramic fashions. These images may not carry enough information for a more global reconstruction of the scene. Here we present state-of-the-art steps we implemented towards solving this problem (Figure 1):

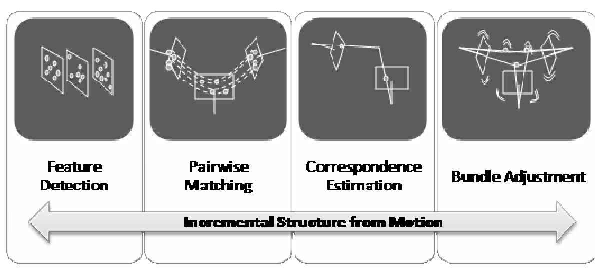


Figure 1. Steps of Structure from Motion from left to right (extended from Nistér and Davidson 2005)

Feature Detection and Correspondence: First step to use site images for reconstruction of the as-built scene is to find distinct features on each image, which allow matching these features across a subset of images. Despite significant research on feature detection and matching, only recently researchers have proposed techniques that prove to be successful in extracting and detecting salient regions (in image/scale space) invariant with respect to scale, orientation and affine transformations. Mikolajczyk et al. (2005) reviews some recently developed view-invariant local image descriptors and experimentally compare their performances.

Structure from Motion (SfM): Aims to reconstruct the unknown 3D structure and estimate camera positions and orientations from a set of image feature correspondences. Photogrammetric techniques such as bundle adjustment (Triggs et al. 1999) are currently used in computer vision for 3D reconstruction and SfM optimization. In our approach, we use the effective method introduced by (Snavely et al. 2007, and, Brown and Lowe 2005). While these techniques have been applied for image navigations and touring, our paper marks the first successful demonstration of SfM technique being applied to geospatial photographs that capture a dynamic construction scene.

Image based Rendering: Following Snavely et al. (2006), image based rendering techniques can be used to synthesize new views of a scene from a set of photographs. In that paper, these techniques are reviewed. Our work is close to Phototour of Snavely et al. (2006) and Sea of Images (Aliaga et al. 2003) where a large set of images is taken throughout architectural spaces. In our work, images are casually acquired on site (as in Snavely et al. 2006), rather than being taken from fixed locations or on guided robot (as in Aliaga et al. 2003).

Discussion on Experimental Setup and Results

Construction photo-logs for our experiment have been collected on two construction sites on a daily basis. One of the authors has been working on these projects and has taken regular daily construction photos. Rather than only taking photos from specific locations or progress within the day, scenes that capture overall depiction of the construction site are captured as well. Figure 2 shows a variety of photo-logs captured for progress analyses in these projects. During the experiment a high-resolution SLR camera was carried. The choice of a high resolution camera was based on the possibility for further enhancement of the algorithm so the quality of the images could be synthetically reduced and the keypoint detection could be tested on synthetically lowered resolution images. To assure the availability of data for further analysis, a larger number of photos than average (about 200/day) have been collected to allow more commonalities between images to be detected. These projects are as follows: (1) Student Dining Hall Project: A two-story masonry and curtain wall LEED Silver certified building with a partial basement of about 139,327sf. This project is a $36M steel frame with composite decking in about 25 months of scheduled work. (2) Residence Hall Project: A two-story masonry and curtain wall LEED Silver certified building with a partial basement of about 58,000sf with structure base being reinforced concrete frame. This project is 21 months and the construction cost is approximately $15M. These projects show two major types of structures which makes them very attractive for our case allowing us to make sure our suggested approach works in different conditions.

Figure 2. An uncalibrated subset of Student Dining/ Residence Hall Photo-log collected for various site analyses using different cameras on a daily basis

The first step is to find feature points in each image that could be used to estimate the initial structure of the scene. In our work, we use the SIFT keypoint detector (Lowe 2004), because of its good invariance to scale changes and view and illumination transformations as well as its standard application in the computer vision domain (Savarese and Fei-Fei 2007). An image of 3MPixels typically gives about 9,000 to 11,000 features. An example of detected features and number of features detected are illustrated in Figures 3 and 4 respectively.

Figure 3. (a) An image taken on 09/27/08 from Student Dining and Residence Hall projects; (b) Same image in grayscale with SIFT feature locations visualized in blue

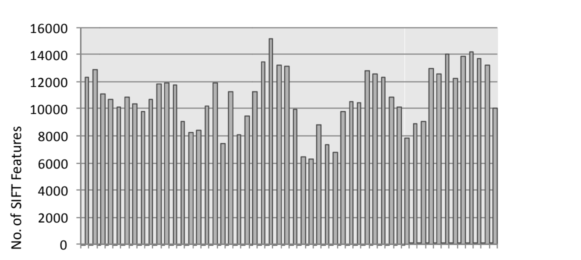


Figure 4. No. of SIFT features on the 52-image subset taken on 09/27/08. No. of SIFT features within the range of [6481, 15160]. Quality of images synthetically reduced to 25% of the original form (Tested image dimensions = 2144 × 1424).

Once the features have been detected over the dataset, we need to detect the number of matching features in each image pair. To minimize computational speed, as experienced by Snavely et al. (2006), we use ANN's priority search algorithm and limit each query to check a limited set. Furthermore, instead of classifying false matches by thresholding the distance to the nearest neighbor, we use the ratio test described by Lowe (2004): for a feature descriptor in image i, we find the two nearest neighbors in j, with distances d1 and d2, then accept the match if d1/d2< 0.6. If more than one feature in i matches the same feature in j, we remove both of such matches, as one of them is a false match. Figure 5-left shows the number of matched SIFT features within the subset. Due to the sensitivity of reconstruction algorithm to false matches, we use an algorithm to remove such false matches. In our approach, once the matching features are detected in an image pair, we robustly estimate a fundamental matrix for the pair using RANSAC (Fischler and Bolles 1981). The fundamental matrix removes false matches as enforces corresponding features to be consistent under viewpoint transformation. In our model, in each iteration of RANSAC, a fundamental matrix is computed using the 8-point algorithm of Hartley and Zisserman (2004), and then the problem is normalized to improve robustness to noises (See Figure 5-right for refined matched features).

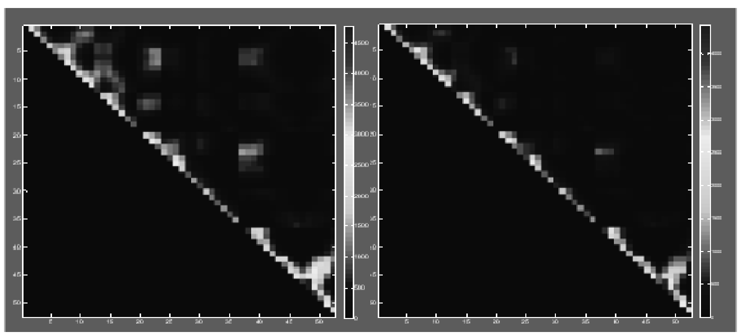


Figure 5. Number of matched SIFT features between each image pair. Both axes show the camera indices and the colored dots visualize the number of SIFT features in image pairs matched before and after the fundamental matrix was fitted to the matching features.

Now, we recover camera extrinsic and intrinsic parameters (extrinsic parameters: rotation, translation; and intrinsic parameters: focal length and distortion) for each image and a 3D location for each keypoint. The recovered parameters should be consistent, in that re-projection error; i.e., sum of distances between the projections of all 3D features and their corresponding image features, is minimized. This minimization problem can be formulated as using the bundle adjustment algorithm (See Triggs et al. 1999 for more details). First, we estimate extrinsic and intrinsic parameters of a single image pair. Since bundle adjustment as other algorithms for solving non-linear problems may get stuck in bad local minima, it is strongly suggested by many researchers (e.g., Snavely et al. 2007, Nistér 2004) to start with a good initial image pair and good estimates for camera parameters in the chosen pair. This initial pair for SfM should have a large number of matches, but also have a large baseline, so that the initial scene can be robustly reconstructed. An image pair that is poorly described by a homographic transformation stratifies this condition. A 2D image homography is a projective transformation that maps points from one image plane to another image plane (Hartley and Zisserman 2004). We find the homography between all image pairs using RANSAC with an outlier threshold of 0.4% of maximum of image width and height, and store the percentage of feature matches that are inliers to the estimated homography. We select the initial image pair as that with the lowest percentage of inliers to the recovered homography, but with at least 100 matches (As experienced by Snavely et al. 2007). The extrinsic camera parameters for this pair are estimated using Nistér's five point algorithm (Nistér 2004), and then the tracks visible in the image pair are triangulated. A two-frame bundle adjustment for this initial pair is performed. Next, we add another camera to the optimization. We choose the camera that examines the largest number of estimated tracks, and initialize the new camera's extrinsic parameters using the Direct Linear Transform (DLT) technique (Hartley and Zisserman 2004) within a RANSAC procedure. For this RANSAC step, we use an outlier threshold of 0.4% of maximum of image width or height. We use focal length from the EXIF - exchangeable image file format- tags of JPEG images to initialize the focal length of the new camera and estimate the intrinsic camera matrix (see Snavely et al. 2007 more details).

Starting from this initial reconstructed scene, we run the bundle adjustment algorithm, allowing new camera and feature points it observes to change while the rest of the model is kept fixed. A feature point is added if it is observed by at least one recovered camera, and if triangulating the location gives a well-conditioned approximation. We estimate the conditioning by considering all pairs of rays that could be used to triangulate that point, and finding the pair of rays with the maximum angle of separation. If this maximum angle is larger than a threshold (As experienced by Snavely et al. 2007) then the point is triangulated. Once the new points have been added, we run another global bundle adjustment to refine the entire as-built reconstructed scene. We use the minimum error solution with the sparse bundle adjustment library of Lourakis and Argyros (2004). This procedure is repeated for all cameras until no remaining camera observes enough reconstructed 3D points to be reliably reconstructed. Overall only a subset of the used images will be reconstructed. This subset is not selected beforehand, but is determined by the algorithm. After the as-built scene is reconstructed, the scene needs to be used for interactive explorations. Authors implemented an image-based rendering system in Microsoft C++ .NET using Microsoft DirectX9 graphics library. The following data structure is used to represent the as-built reconstructed scene: (1) A set of keypoints, in which each keypoint consists of a 3D location and a color that is averaged out from all the progress images that the keypoint is being observed from; (2) A set of cameras, while the extrinsic parameters (translation and rotation), and intrinsic parameters (focal length and distortion in height and width directions) are estimated; and (3) A mapping between each point and all the cameras that observe the point. A list of numbers of cameras which observe the point, the location of the point in local coordinates of the image, and the SIFT keypoint index are all stored. While this information is stored, cameras would be rendered as frusta. Figure 6-a & b show the reconstructed sparse scene from the same image subset and illustrate 6 of the registered cameras. Once a camera is visited in this reconstructed scene, the camera frustum is texture-mapped with a full resolution of the image so user can zoom in and acquire progress and productivity details as well as workspace logistics. Figure 6 – c, d, e and f show the location of a frustum textured while demonstrating how the site image is geo-registered with the as-built point cloud.

Figure 6. Sparsely reconstructed scene of Residence Hall using 52 images with 25% of image qualities. Six camera frusta are rendered and geo-registered.

Conclusion

Our proposed system marks the first system that allows as-built construction spatial information to be visualized using unordered site photo-logs. The demonstrated system has the following benefits: (1) *Remote Construction Control Decision Making*: It allows project managers, superintendents and other project participants to virtually walk on the construction site, as-of the time the scene has been reconstructed and position themselves in those positions that progress images have been taken. Such an interactive user walk-through allows progress to be discussed. (2) It *minimizes the time required to discuss the as-built scene*: Project managers and superintendents will spend less amount of time discussing or explaining progress. Rather, they can spend more time on how a control decision could be made. Furthermore, reconstructed as-built scene and geo-registered images allow workspace logistics, safety issues, progress and even productivity of workforce and machinery to be remotely analyzed. Such an as-built system could also be beneficial in weekly contractor coordination meetings, as the workspace could be navigated through the virtual world. (3) *Significant cuts in travel time and cost on project executives and architects* – Project Executives and architects can study the reconstructed scene and geo-registered images, instead of spending time and money traveling to the jobsite. The reconstructed scene with as-built progress images can be beneficial, especially when the possibility of adding new photographs quickly to the system is considered. Even if a vanishing point of an interest is not registered within the reconstructed scene and is not present in geo-registered image dataset, the user (e.g., owner, project executive) can request the scene to be photographed, since the geo-registration removes confusion on perspective which is inherent in dynamic scenes. Those photographs taken can also be quickly geo-registered which allows a significant problem of progress communication to be resolved. (4) *D^4AR System- 4 Dimensional Augmented Reality Tool* - This system could also be used as a baseline for an Augmented Reality tool wherein as-planned model could be geo-registered within the spatial as-built environment allowing construction progress deviations to be measured, analyzed and communicated. To that extent, authors have proposed D4AR – 4 Dimensional Augmented Reality - system which superimposes the planned model over point cloud and utilizes a traffic light color spectrum for visualizing progress (Golparvar-Fard et al. 2009); (5) *Automatic progress tracking*- Since this model geo-registers construction site photographs, it could serve as a rich baseline for automating progress monitoring through consistent visual detection of progress deviation and comparison with as-planned information; (6) *Registering New Daily Site Photographs*- New construction progress photographs can be registered within the system instantly. First, the user can open a set of progress images, and drag and drop each image onto its approximate location on the as-planned model. After each image has been dropped, the system can estimate location, orientation, and focal length of the new image by running the SfM algorithm. First, SIFT keypoints are extracted and matched to the keypoints of the cameras closest to the initial location; then the existing 3D points corresponding to the matches are identified; and finally, these matches are used to refine the pose of the new camera. Our preliminary results show perceived benefits and future potential enhancement of this new technology in construction, in all fronts of automatic data collection, processing and communication; though there are still many technical challenges in developing a full systematic progress monitoring system. These are currently being explored within the research projects highlighted in this paper.

Acknowledgements

We would like to thank Turner, Williams Brothers, Grunloh Construction companies especially Mr. Greg Cuttell, Mr. Bob Bursack and Mr. Nick Canellis from Turner as well as University of Illinois- Housing, Facilities and Services for their contributions to this project. This work is financially supported by NSF Award No. CMMI-0800500.

References

[1] Abeid J., Allouche E., Arditi D., and Hayman M. (2003). PHOTO-NET II: a Computer-based Monitoring System Applied to Project Management, J of Automation in Constr, 12 (5), 603-616.

[2] Aliaga D., Funkhouser T., Yanovsky D., and Carlbom I. (2003). Sea of images. IEEE Computer Graphics and Applications, 23 (6), 22-30.

[3] El-Omari S. and Moselhi O. (2008). "Integrating 3D laser scanning and photogrammetry for progress measurement of construction work". J. of Automation in Const. 18 (1), 1-9.

[4] Golparvar-Fard M., Savarese S. and Peña-Mora F. (2009). "Interactive visualization of construction progress monitoring with D4AR – 4 Dimensional Augmented Reality- models." Proc., Construction Research Congress, Seattle, WA, Apr 2009.

[5] Golparvar Fard, M., Sridharan, A., Lee, S. and Peña-Mora, F. (2007) "Visualization of construction progress monitoring on time-lapsed photographs", Proc., Constr. Mgmt & Econ, Reading, UK

[6] Golparvar Fard M., Staub-French S., Po B. and Tory M. (2006). "Requirements for a mobile interactive workspace to support design development and coordination." Proc., XI Joint Int. Conf. on Computing in Civil & Bldg Eng ASCE, Montreal, QC, 3587-3596.
[7] Hartley R. and Zisserman A. (2004). Multiple view geometry, Cambridge Univ. Press, UK.
[8] Jaselskis E., Cackler E., Walters R., Zhang J., and Kaewmoracharoen M. (2006). "Using scanning lasers for real-time pavement thickness measurement". CTRE Project 05-205, National Concrete Pavement Tech Center, Iowa State Univ.
[9] Kiziltas S., Akinci B., Ergen E. and Tang, P.o (2008). "Technological assessment and process implications of field data capture technologies for construction and facility/infrastructure management", ITcon, 13, Sp. Issue Sensors in Const. and Infrastructure Mgmt., 134-154.
[10] Lourakis M. and Argyros A. (2004). "The design and implementation of a generic sparse bundle adjustment software package based on the Levenberg–Marquardt algorithm" (Technical Report 340). Inst. of Computer Science-FORTH, Heraklion, Crete, Greece.
[11] Lowe D. (2004). "Distinctive image features from scale-invariant keypoints". Int Journal of Computer Vision, Vol. 60, No. 2, 91-110.
[12] Meredith J. and Mantel S. (2003). Project Mgmt: A managerial approach. J. Wiley & Sons, 5th Ed.
[13] Mikolajczyk K., Tuytelaars T., Schmid C., Zisserman A., Matas J., Schaffalitzky F., Kadir T., and van Gool L. (2005). "A comparison of affine region detectors", Int J. of Computer Vis., 65 (½), 43-72.
[14] Navon R. and Sacks R. (2007). "Assessing research in automated project performance control (APPC)." J. of Automation in Construction, 16 (4), 474-484.
[15] Nistér D. (2004). "An efficient solution to the five-point relative pose problem", IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 26 (6), 756-770.
[16] Nistér D. and Davison A. (2005). "Real-Time motion and structure estimation from moving cameras", Tutorial at CVRP 2005.
[17] Oglesby C., Parker H., Howell G. (1989). Productivity improvement in construction, McGraw-Hill.
[18] Poku S. and Arditi D. (2006). "Construction scheduling and progress control using geographical information systems". J. of Computing in Civil Eng., 20 (5), 351-360.
[19] Savarese S. and Fei-Fei L. (2007). :3D generic object categorization, localization and pose estimation". IEEE Intern. Conf. in Computer Vision (ICCV). - Brazil, October, 2007.
[20] Shih N.J., Wu M.C., and Kunz J. (2004). "The inspections of as-built construction records by 3D point clouds". Stanford Univ., CIFE WP #090.
[21] Snavely N., Seitz S., and Szeliski R. (2006). "Photo tourism: exploring photo collections in 3D." Proc., ACM Transactions on Graphics, 25 (3), 835-846.
[22] Snavely N., Seitz S., and Szeliski R. (2007). "Modeling the world from internet photo collections". Int. J. of Computer Vision.
[23] Snavely N., Seitz S., and Szeliski R. (2008). "Finding path through the world's photos". Proc., ACM Transactions on Graphics, 27 (3), 11-21.
[24] Soibelman L., Wu J., Caldas C., Brilakis I. and Lin K.Y. (2008). "Management and analysis of unstructured construction data types." J. of Advanced Engineering Informatics, 22, 15–27.
[25] Su Y., Hashash Y., and Liu L.Y. (2006). "Integration of construction as-built data via laser scanning with geotechnical monitoring of urban excavation". J. of Const Eng. Mgmt, 132 (12), 1234-1241.
[26] Teizer J., Kim C., Haas C., Liapi K., and Caldas C. (2005). "Framework for real-time three-dimensional modelling of infrastructure". Geology and Properties of Earth Materials 2005, Transportation Research Board Natl Research Council, Washington, 177-186.

Rich Knowledge Parametric Tools for Concrete Masonry Design Automation of Preliminary Structural Analysis, Detailing and Specifications

Andres Cavieres[1] , Russell Gentry[2] and Tristan Al-Haddad[3]

[1]Georgia Institute of Technology, College of Architecture, Atlanta, Georgia, United States, andres.cavieres@gatech.edu
[2]Georgia Institute of Technology, College of Architecture, Atlanta, Georgia, United States, T.Gentry@coa.gatech.edu
[3]Georgia Institute of Technology, College of Architecture, Atlanta, Georgia, United States, tristan.alhaddad@coa.gatech.edu

Abstract

Our research focuses on how new computational tools and design methods can support the design process of masonry buildings. Specifically we discuss the potential of parametric modeling technologies to promote innovation by embedding knowledge on masonry construction as both generative rules and checker functions. Their goal is to inform designers on the feasibility of their intents on early stages of the design process. For this purpose we will adopt a methodology and notation called Building Object Behavior (BOB) to identify and translate construction knowledge towards the implementation of masonry parametric objects. A curved masonry wall is used as case study for the development of several parametric prototypes and their possible implications for collaboration are discussed.

Keywords: BIM, Knowledge-Based Design; Parametric Modeling, Masonry Building Design

Some Difficulties in Masonry Design

Research in masonry construction have promoted important innovations on new masonry unit types, structural analysis methods and more efficient construction processes for years (Beall 2000; Beall & Jaffe 2003; Ramamurthy & Kunhanandan 2004). However there is still a perception that the limits of masonry design are not being challenged by the architects. Current technical innovations are not being extensively transferred into architectural design practice and in most cases new masonry buildings continue to adopt conventional and rather conservative solutions.

There are several causes for this problem but two of them seem to be the key limitations that are currently threatening the competitiveness of this material. The first relates with an increasing number of misconceptions and prejudices among architects regarding the high cost of masonry construction and its limited formal possibilities. Such misconceptions are mainly due the lack of knowledge, especially among young architects about different masonry types and their architectural possibilities.

It can be argued that challenging masonry limits requires a considerable level of expertise, which is in contradiction with the lack of knowledge. Furthermore, expertise on masonry design is, as in any other type of domain, rare and expensive. Construction of complex or unusual configurations is naturally considered risky and complicated without expert knowledge and judgment. To avoid this problem the easy choice is the adoption of conventional solutions or simple variations of well known typologies that reduce risks and uncertainties. In this manner possible difficulties are kept under control by adopting prescribed design formulas and specifications.

A second relevant reason besides the above mentioned lack of technical knowledge is the reduced number of design tools currently available to represent and explore new masonry configurations in a more efficient and productive way. In the absence of such tools what becomes costly is the amount of effort architects have to put on modeling and detailing a building that could have hundreds or probably thousands of masonry units.

Our research aims to address this issue focusing on the development of rich-knowledge parametric modeling tools under a Building Information Modeling (BIM) framework. For that purpose we explore some methodologies and examples developed for the structural steel and pre-cast concrete industries to capture and embed their domain knowledge into parametric modeling tools. The final goal is to improve the design and construction processes by supporting the creation, testing and evaluation of a higher number of design alternatives from the beginning.

The Limitations of Current CAD Technologies

Masonry construction implies the placement and bonding of individual modular units into a continuum. The geometric characteristics of this continuum as well as the specific patterns adopted for the placement of units are part of a sequence of decisions made by architects and engineers while trying to solve a set of simultaneous issues. Structural stability, functionality, aesthetics and satisfactory performance are some of the fundamental requirements that any building must satisfy as part of a design problem. To provide an integrated solution is not easy, and whenever a solution seems to be reached the chances for exploring other valid alternatives are going to be reduced.

However, good solutions often emerge from exhaustive comparison of different models from a wider pool of alternatives. Unfortunately designers usually succumb to the temptation of believing that the first options could be good enough; or they simply stick to what they already know, avoiding in this manner further explorations.

This phenomenon of accepting the "already-known" or the quickest solution is known "design fixation" (Purcell & Gero 1996; Buelow 2007) and we argue that the tendency to follow it becomes particularly strong when the number of parts involved in a problem is high. A continuum assembled by several small units, i.e., a masonry wall, a cladding system, a tiled roof or a brick pavement arise as typical situations where the number of parts implies such number of interdependencies that any further exploration beyond what is commonly accepted is taken as an extra-effort. The fixation on conventional configurations is generally accepted as "default" and in some cases it is even justified under claiming of adherence to traditional constructive wisdom.

This over-simplification is evident on conventions architects adopt for representation of masonry elements in CAD systems. The complexity of a masonry wall is reduced by denoting it as a homogeneous volume, i.e., a rectangular block that does not describe the geometry of the composing units nor their bonding pattern. In terms of quantification of units, a simple calculation is made as function of the total area to be covered, the size of the chosen unit and the type of mortar required. In this manner CAD systems are not used to geometrically represent all the components nor consider special cases such as block cuts or custom units usually required in any masonry design.

If a designer intends to explore different types of building forms, unusual bonding patterns or special details, a complete representation of the masonry assembly would be the most appropriate way to get better visualization and understanding. In a conventional CAD system this requirement would force him / her to model all the units one by one, place them on their respective positions and set all the joints, angles and articulations "by hand". Once the assembly is modeled any modification on design implies an excessively time consuming and error prone manual process that would easily discourage most attempts.

Knowledge Based Parametric Design

There are currently important software development and research being done to solve similar issues of geometric and information complexity. Their first goal is to take advantage of computing power to automate repetitive tasks such as production of specifications and detailing. These developments strongly suggest that traditional CAD methods that represent building assemblies by oversimplifying their geometry are going to be replaced by another one where all relevant parts will be explicitly modeled. The potential benefit of such type of representation lies on the intelligent behavior that parametric components can exhibit.

Indeed this intelligent, knowledge-based behavior is the most promising capability of parametric technologies in design. The reason is that parametric modeling provides mechanisms to embed domain expertise and design intent into a set of topological and geometric relationships, allowing not only automation of low-level tasks but most importantly, to assist and inform designers with relevant knowledge about the validity of current design alternatives (Eastman, Lee & Sacks 2003).

Design knowledge usually is expressed in terms of some system of rules that control the form. According to Robert Aish (Aish 2005) design is fundamentally about the creation of these rules by means of the definition of proper relationships between design components. Many of these relationships are geometric in nature or can be expressed in geometric terms, containing both knowledge and intentionality that designers have regarding a given problem.

The relation between geometry and knowledge however is not specific of parametric models but property of any design representation. The distinction here lies on the fact that in a conventional representation, such as paper drawings or non parametric CAD models, the knowledge is normally tacit, hidden behind the geometry while the rules that define relationships only exist in the mind of the designer (Katz 2007). In a parametric model instead this order is inverted; the system of rules becomes explicitly and easily available while the geometry represents just an instance of it.

A practical consequence of this inversion is that the rules and constraints that control a geometric form can come from technical requirements of other domains, such as structural analysis, fabrication or construction management, which usually has a secondary or late participation in decision making. It is anticipated that bringing these requirements in the form of knowledge early on conceptual stages of the design process, more feasible and innovative solutions could be explored from the beginning. This is important because is at this level where most sensitive and costly decisions are made in building design (Eastman et al. 2008).

In this manner parametric design offers the possibility of generation of multiple variations and design alternatives that can be produced, optimized, tested and selected systematically according to various methods. Decision making then has the potential to become not only a more systematic and objective process, but also a much more collaborative process due the availability of explicit, readable and shared knowledge.

Upon that the ability to automatically access and manipulate geometric dependent data, making it interoperable and reusable is the key feature of parametric systems that allows the generation of rich building information. A fully detailed 3D model becomes then the base representation for building projects, working as a central data source for architects, engineers and contractors during the entire life cycle of the project (Lee, G., Sacks, R. & Eastman, C. 2005).

Building Object Behavior

According to Eastman one of the major challenges in the development of such rich-knowledge parametric systems is to find a general method to facilitate the translation of relevant expertise into a set of functional parametric objects. To solve this issue the Building Object Behavior (BOB) description method and notation was developed by his team to assist in this task during the specification of a parametric modeling system for the North American precast concrete industry (Sacks, R., Eastman, C. & Lee, G. 2003). Nevertheless BOB was developed not only focusing on the precast concrete industry, but rather though to be useful on other types of design problems and construction systems. In this manner the general and abstract nature of BOB methodology and notation allowed us to adopt it as basic reference for the definition of parametric behavior of masonry models.

Object behavior in a parametric modeling environment is seen as the ability of a building component or assembly to respond to an internal or external stimulus preserving the original design intent. This response occurs when the system is capable of taking automatic actions in order to "maintain the topological and geometrical consistency of the relationships within and between model objects". In this manner objects have to be modeled not only as they look but most importantly, as semantic relationships within a specific domain (Sacks et al. 2003).

However a major issue for the implementation of domain specific BIM parametric solutions relates with the problem of how to specify and embed relevant design and engineering knowledge in a parametric modeling system. One of the main difficulties arises from the fact that much of this knowledge is tacit and very hard to be explicitly represented. The tacit nature of design knowledge usually produces different interpretations at the implementation level that may cause several implications downstream the process, such as ambiguity and idiosyncratic adaptation of models.

The fact that a same object behavior can be implemented in different ways raises the question about efficiency, re-usability and scalability of the solution. Therefore there was an urgent need for a method that could help to rapidly and abstractly capture and represent such behaviors prior any software implementation or modeling activity. This method should help to pre-tune and guide parametric object definition in a testing phase, in such a manner that ambiguity and unnecessary complexity could be reduced. According to Lee (Lee

2005) this approach was also found to be especially appropriate for relatively large and complex parametric modeling activities where effective collaboration is a critical issue.

BOB Specification

The building object behavior (BOB) notation was developed to help designers to define a set of parameters and the relationships that represents the main aspects of a building object and its expected behavior within a given domain. According to Lee it is essentially graphical shorthand for sharing the descriptions among different members of a collaborative setting.

In this notation each shape can be labeled as a real-world object with a unique identifier (e.g., room A, wall W1, a column and so on). The graphical conventions adopted by the notation are the same of engineering drawing; however the drawings do not intend to be an accurate representation of shapes. Only the topological structures are important since the goal is to capture and clearly communicate relevant relationships existing within or between objects. All the accuracy will be parametrically determined by design afterwards (Figure 1).

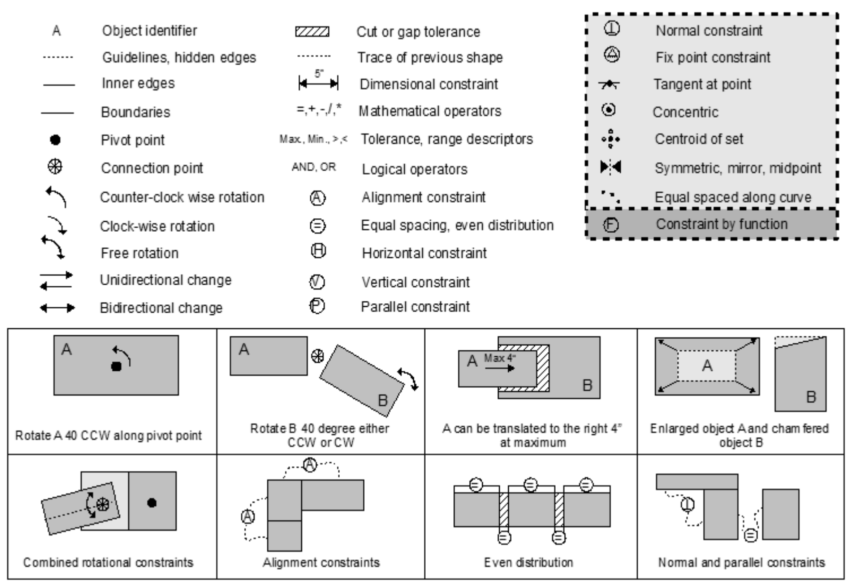


Figure 1: Set of primitive constraint types and extended set declarative geometric constraints. White rectangle set extracted from Lee et al.(2006). Clear grey extracted or based on Bettig (2003). The lower diagrams show examples of parametric behaviors declared using BOB notation as defined by Lee (2006).

In this research the implementation of knowledge of masonry design was done through the development of parametric prototypes using an existing parametric CAD system, called GenerativeComponentstm | [1]. Curved walls were adopted as case study to be defined parametrically using BOB, in such a way that solutions can be easily extended towards more conventional flat orthogonal walls. So far this process has followed three out of four base steps considered by BOB: 1) Elucidation, 2) Translation and 3) Implementation. The fourth point, validation is pending until a more extensive set of implement models are produced.

Elucidation: A curved concrete block masonry wall was chosen as case study for the definition of parametric behavior using BOB. The election of a curved geometry allowed the implementation of general solutions that can be easily adapted towards more common flat orthogonal walls. The first step describes the expected behavior of both a single concrete masonry unit (CMU) and an entire wall assembly according to

1 DigitalProject and Revit are being considered for future developments.

masonry construction guidelines. For this purpose we followed the technical specifications and design recommendations in a series of technical detailing documents provided by the National Concrete Masonry Association of North America (NCMA, 2002, 2003, 2004, 2005, 2006).

The elucidation process starts by identifying the most relevant relationships exiting in a design problem. Such relationships do not need to be explicitly geometric but at least have to provide the basis for a geometric interpretation. Therefore the potential link between design knowledge and its parametric representation lies on the clear identification of the building components that might change and how they should do so.

Translation: At the most fundamental level the geometrical representation of domain knowledge operates as relationships between elemental components of the Euclidean geometry, that is, points, lines and surfaces. The set of primitive constraint types and constraint declarations showed above have to be applied to a simplified representation of the of a CMU unit, and after that to the entire curved wall assembly.

Figure 2 shows the sequence of geometric and declarative constraints that defines the behavior of a single unit. The unit itself is treated as a hierarchical assembly of points, lines and surfaces. Upon these basic elements higher level geometric entities are built.

In higher level the wall assembly is defined following a top-down modeling approach. That means that the overall skeleton or control rig is created before the propagation of the above specified CMU unit. The initial input is a surface that describes that wall geometry. This surface works as supporting element for the wall assembly and at this level the bonding patterns and the spacing between courses, headers and mortar thickness is defined.

For this purpose horizontal section curves are specified (controlled by horizontal constraint). The distance between curves should be the height of the CMU head plus a given mortar thickness (equal spacing constraint). Once the course curves are specified a two-dimensional array of points has to be defined by propagating points along each curve at constant spacing (equal spacing along curve).

The propagated points work as reference for the propagation of vertical segments (vertical constraint) which represent a two-dimensional array of masonry headings for a doubled-curved wall based on corbeling courses. Each pair of these vertical segments provides the basic four input points (two start points and two end points) for the propagation of rectangular polygons that work as place holder for the final insertion of CMU blocks (see previous section).

However that positioning of the place holders should not be even but rather follow a specified bonding pattern type. This requirement is accomplished by the specification of a higher level function that combines several geometric and declarative constraints within an algorithm. Therefore a special function based on iterations has to be defined at the implementation level in order to create such bonding pattern.

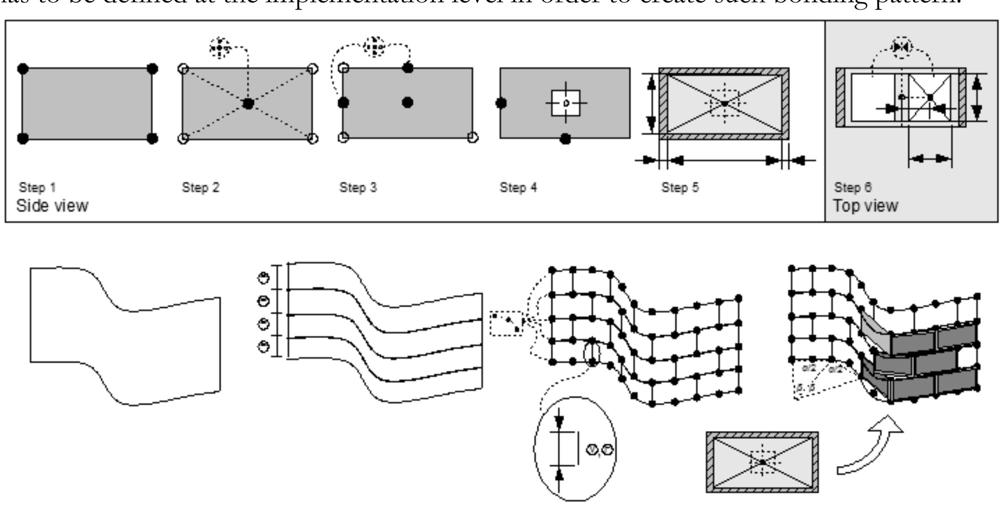


Figure 2 Sequence of aggregated parametric behaviors from CMU blocks to a block wall assembly using BOB notation.

Implementation: Once the basic rules regarding masonry bonding patterns and corbeling curved wall were established using BOB notation, the implementation of parametric building objects was grandly facilitated. However some detailed aspects of the translation had to be solved in combination with the implementation process.

For the running bond pattern a simple algorithm was defined and then implemented as an iteration to create the woven sequence. For the generation of rebar a similar function was created, but also allowing extra input for structural specification of spacing and diameters of each rebar according to load calculations (Figure 3). The generative nature of these scripts allows the designer to freely change the dimensions and shape of the supporting surface while keeping the regularity of the running bond pattern and the specified rebar spacing.

At this point the problem about the structural and constructive feasibility of a non- conventional geometry arose as a main issue. The limits of wall curvature and the inclination of the wall by corbeling consecutive courses were taken as exemplary cases for implementation of structural and construction rules.

Although these rules can be embedded a priori in the same way as generative rules, so that only valid forms can be created by 'enforcement', we considered it an important limitation for free design exploration. Therefore it was foreseen that verification a posteriori by adopting checking functions would solve in a better way both requirements of free exploration and constructive feasibility. The automatics enforcement would be avoided by relying on the fact that the system can inform were undesirable limits are being reached and by letting the design team to decide whether such limits are acceptable or not.

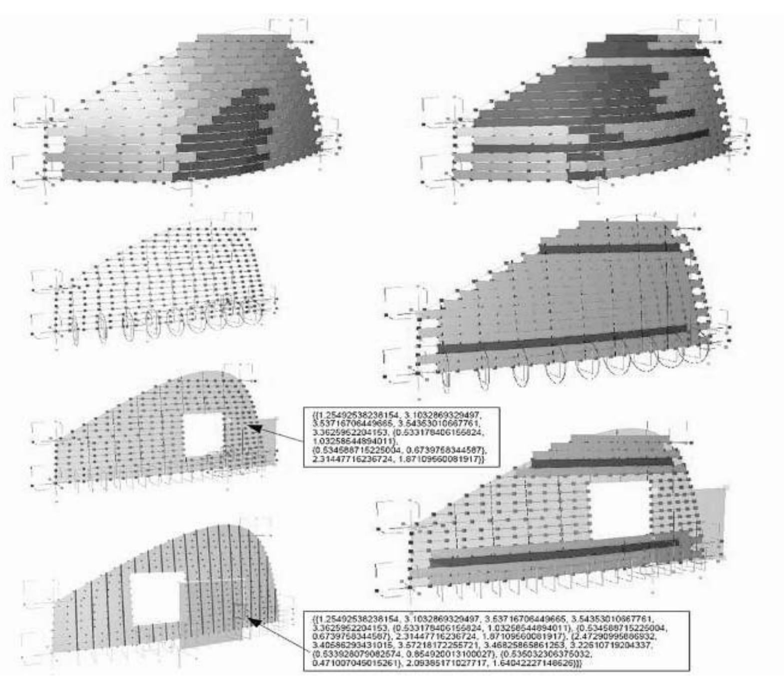


Figure 3: Examples of generative rules and checker functions. Corbelling and angles between adjacent units are calculated on demand by activation of special corbelling or angle checker functions. Masonry running bond pattern and vertical reinforcement are generated based on predefined patterns and automatic calculations of load and stresses.

Validation: Currently most of our research effort is being put on the three initial stages required for parametric definition of masonry objects. The validity of the prototypes done so far is currently being tested by the authors by including them as part of small scale design exercises. The goal is to have several working prototypes that include more design and construction rules from the NCMA technical specifications and guidelines. The second goal is to generate effective interfaces for collaboration with other specialists, specifically for structural analysis based on spreadsheets or customized interoperability with domain specific analysis tools. Those interfaces are intended for design refinement and optimization, by taking advantage of commonly specified object's behaviors.

Once a comprehensive set of design rules and guidelines are implemented then more extensive design exercises will be done to test the effectiveness of the prototypes under a real-world design process work flow. The evaluation will look that the behavior displayed by parametric assemblies reflects the knowledge and design intent embedded on them at the beginning by producing the appropriate response and output.

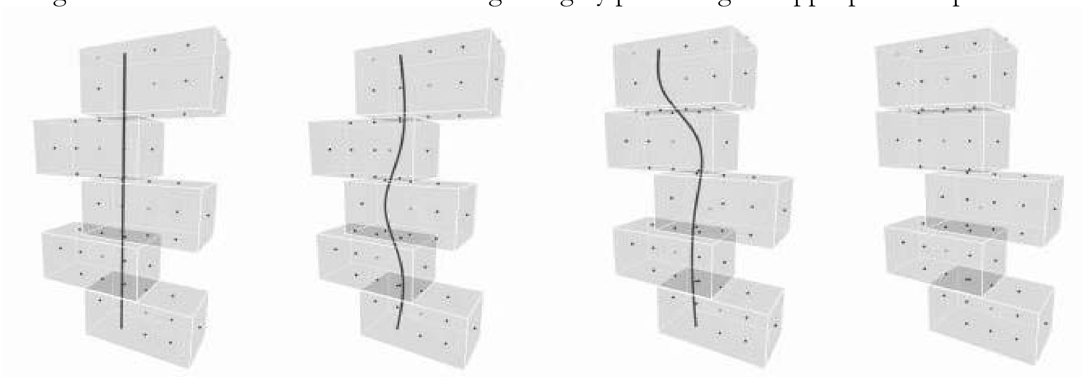
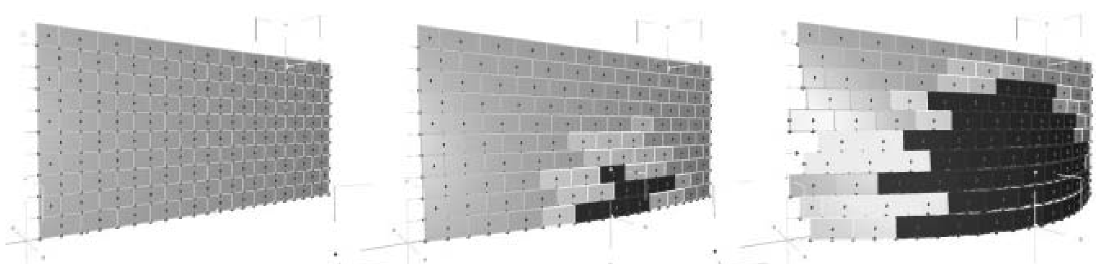


Figure 4: Example of checker function (in development) to determine maximum bending for vertical reinforcement according to wall curvature.



	column01	column02	column03	column04	column05	column06	column07	column08	column09	column10
course01	180	180	180	180	179	179	180	180	180	
course02	180	179	179	179	178	178	178	179	179	180
course03	179	179	178	177	177	177	178	178	179	
course04	179	178	178	177	176	175	176	177	178	179
course05	179	178	176	175	174	175	176	177	178	179
course06	179	178	177	175	174	173	174	176	177	178
course07	178	177	175	174	172	172	174	176	177	178
course08	179	177	176	174	172	170	172	174	176	177
course09	178	177	175	172	170	169	172	174	176	178
course10	179	177	176	173	170	168	169	172	175	177

	column01	column02	column03
ecc01	0.00	1.12	2.17
ecc02	0.55	1.62	2.60
ecc03	0.00	1.06	2.07
ecc04	0.51	1.51	2.44
ecc05	0.00	0.95	1.88
ecc06	0.45	1.33	2.20
ecc07	0.00	0.83	1.65
ecc08	0.38	1.14	1.91
ecc09	0.00	0.69	1.40

	column01	column02	column03
course01	286.25	286.25	286.25
course02	326.25	326.25	326.25
course03	366.25	366.25	366.25
course04	406.25	406.25	406.25
course05	446.25	446.25	446.25
course06	486.25	486.25	486.25
course07	526.25	526.25	526.25
course08	566.25	566.25	566.25
course09	606.25	606.25	606.25

	column01	column02	column03
course01	7.15625	7.15625	7.15625
course02	7.231367	5.40928	3.670329
course03	9.15625	5.045189	1.054585
course04	6.755317	0.049048	-6.36895
course05	11.15625	1.344295	-8.20291
course06	5.546491	-7.50564	-20.0501
course07	13.15625	-3.682	-20.1179
course08	3.749832	-16.8342	-36.7195
course09	15.15625	-9.73739	-34.1217

Figure 5: Example of function to check angles between adjacent CMU blocks in a curving wall. Grey block means angle values above 179 degrees, so that no cut is necessary according to guidelines. Yellow means that flanges of flanged blocks have to be removed. Red means cuts beyond flanges. A spreadsheet file of angle values is automatically generated (top), as well as spreadsheets containing eccentricities between adjacent rows (bottom left) and automatic load and stresses calculations for vertical reinforcement (bottom right). The hierarchical structure of the parametric assembly can be represented by a Directed Acyclic Graph (DAG) that manages all changes by automatically propagating updated values to dependent components.

Summary and Future Works

Parametric modeling is a core technology of Building Information Models (BIM) that enable the generation of rich building information. This technology promotes more accurate representations of building components and assemblies, facilitating the modeling of complex systems and automating tedious and error prone activities of detailing, scheduling and control of changes. Nevertheless the most important feature of parametric modeling is that it supports the incorporation of technical knowledge on early stages of the design process. In this manner we believe that these tools have the potential to improve design processes by increasing designer's understanding on the implications of his / her decisions and promoting the collaborative exploration of more innovative solutions.

Our research means a step towards such direction, by adapting the Building Object Behavior methodology and notation originally developed for the concrete precast industry into the domain of concrete masonry design.

In an initial stage we started the elucidation of knowledge based on available masonry design guidelines provided by the National Concrete Masonry Association. These guidelines contain both design standards as well as construction best practice. Thus initial samples of construction and structural knowledge were translated into parametric rules using BOB. Several prototypes were implemented including generative rules such as running bond pattern and rebar distribution while the verification of allowable angles between units and corbeling were included as checker functions to be applied a posteriori. This approach was considered useful to avoid the risk of over-constraining and interference to free exploration of alternatives. An important issue will be the development of criteria to define where certain type of rules should be implemented in a generative way or as checker functions, especially considering the implication that they can produce in a design work flow.

Further work has to focus on increasing the number of construction rules to be embedded in both CMU units as well as overall building assemblies (besides walls, masonry pavements, slabs and roofs can be considered).

While the partial prototypes developed so far worked well as independent models, extensive validation tests have to be done under the conditions of complete design processes. This challenge raises the problem of scalability of the model according to computer memory requirements.

Finally, consistent interoperability between design parametric models and engineering analysis tools have to be studied in order to streamline the feedback cycles between engineers and architects. In this scenario it is foreseen that BOB notation and methodology can be an important resource for the implementation of models oriented towards the optimization of many building life-cycle aspects, including fabrication, construction coordination, energy performance, and compliance with performance based building codes among others.

References

[1] Aish, R. “Notes about Generative Components.” SmartGeometry – 23rd Conference of the Education and Research in Computer Aided Architectural Design in Europe ECAADE. 2005. Lisbon, Portugal: Jose Duarte.

[2] Beall, C. (2000). “New Masonry Products and Materials.” Prog. Struct. Engng Mater.

[3] Beall, C. & Jaffe, R. (2002). “Concrete and Masonry Databook”. New York: McGraw Hill.

[4] Bettig, B. & Shah, J. (2003). “Derivation of a Standard set of Geometric Constraints for Parametric Modeling and Data Exchange.” Computer-Aided Design, 33, 17-33.

[5] Buelow, P.V. (2007). “An Intelligent Genetic Design Tool (IGDT). Applied to the Exploration of Architectural Trussed Structural Systems, in Structural Engineering.” Stuttgart, Germany. Universität Stuttgart, Institute of Lightweight Structures and Construction.

[6] Ramamurthy, K., & Kunhanandan Nambiar E., K., (2004). “Accelerated Masonry Construction: Eeview and Future Prospects,.” in Prog. Struct. Engng Mater.

[7] Eastman, C., (2004). New Methods of Architecture and Building. Proceedings of the 24th Conference of the Association for Computer Aided Design in Architecture Acadia. Cambridge and Toronto, Ontario, Canada: S. Williamson, P. Beesley, N. Cheng.

[8] Eastman, C., Teicholz, P., Sacks, R. & Liston, K. (2008). “BIM Handbook: A Guide to Building Modeling for Owners, Managers, Designers, Engineers and Contractors.” Hoboken, NJ : Wiley.

[9] Eastman, C., Lee G. &Sacks, R. (2003). "Development of a Knowledge-Rich CAD System for the North American Precast Concrete Industry." Proceeding of the 23rd Conference of the Association for Computer Aided Design in Architecture Acadia (pp. 16). Muncie, IN: K. Klinger.
[10] Katz, N. (2007). "Parametric Modeling in AutoCAD." (http://aecbytes.com/viewpoint/2007/issue_32.html). AECbytes: Lachmi Khemlani.
[11] Lee, G., Sacks, R. & Eastman, C. (2005). "Specifying Parametric Building Object Behavior (BOB) for a Building Information Modeling System." Automation in Construction, 15, 758 – 776.
[12] Purcell T., & Gero J. (1996). "Design and Other Types of Fixation." Design Studies, 17, 363-383.
[13] Sacks, R., Eastman, C. & Lee, G. (2003). "Parametric 3D Modeling in Building Construction with Examples from Precast Concrete." Automation in Construction, 13, 291– 312.
[14] National Concrete Masonry Association. (2002). "Typical Sizes and Shapes of Concrete Masonry Units." TEK 2-1A Unit Properties. Herndon, VA.: Cemex.
[15] National Concrete Masonry Association. (2005). "Grouting Concrete Masonry Walls." TEK 3-2A Construction. Herndon, VA: Cemex.
[16] National Concrete Masonry Association. (2004). "Modular Layout of Concrete Masonry." TEK 5-12 Details. Herndon, VA.: Cemex.
[17] National Concrete Masonry Association. (2006). "Concrete Masonry Radial Wall Details." TEK 5-10A Details. Herndon, VA: Cemex.

Framework to Improve Mobile Robot's Navigation Using Wireless Sensor Modules

Diego Martinez [1] and Yong K. Cho [2]

[1]Graduate Research Assistant, Computer Engineering and Electronics Department, College of Engineering, University of Nebraska-Lincoln, 1110 S. 67th St. Peter Kiewit Institute (PKI) 118A, Omaha, NE 69182; PH(402)554-3277;FAX (402)554- 3850; email: dmartinez@mail.unomaha.edu
[2]Assistant Professor, Construction Engineering and Management Division, Charles Durham School of Architectural Engineering and Construction, College of Engineering, University of Nebraska-Lincoln, 1110 S. 67th St. Peter Kiewit Institute (PKI) 104C, Omaha, NE 69182; PH(402)554-3277;FAX (402)554- 3850; email: ycho2@unl.edu

Abstract

As on-going research, this paper presents a framework to improve wireless mobile robot's navigational accuracy in diverse indoor environments where the signals are affected by various types of interference including electromagnetic, multi-path, and fading and scattering. In particular, indoor construction environments pose unique challenges to accurate wireless navigation due to their relative complexity and inherently dynamic nature. Several integrated location and orientation sensors including a digital compass, a gyroscope, wheel encoders, an accelerometer, and Ultra Wideband (UWB) position tracking sensors are introduced in this paper. A distinct cause of error for each sensor is studied based on location, traveling distance, and rotational angle. To improve the position data accuracy, statistical methods such as outlier analysis and the Kalman Filter are applied in this research. A framework for position and orientation error compensation between relative and absolute sensors is described with preliminary research results indicating that position and orientation errors can be statistically adjusted in real time.

Keywords: mobile robot, navigation, dead reckoning, kalman filter, wireless sensor, error

1 Introduction

The competitive, market oriented, and rationalized construction of tomorrow will require developing automated and robotized construction system today [4]. This includes indoor construction applications such as interior finishing, piping, excavation, mining, and earth moving [5] among others. In particular, implementing suitable indoor localization in construction processes will lead to an increase in productivity and improvement in work quality and working conditions [4].

The requirement to have reliable positioning is becoming increasingly important, and can be used for indoor position application such as implementing robots with mobile platform for construction tasks [4]. However, implementing mobile robots for construction tasks has proven to be difficult due to the dynamic and uncertain nature of the construction site [5].

It is critical that mobile robot's absolute and relative positions are accurately determined in both outdoor and indoor environments. Global Positioning Systems (GPS) are widely used to determine the absolute position in outdoor environments where the signals are not obstructed by nearby buildings or trees. It is possible to determine the mobile robot's position with an accuracy of 2-3 cm with well-equipped GPS systems. However, GPS signals cannot travel through walls and thus suffer from signal attenuation, making GPS systems unsuitable for indoor applications. One of the most prominent technologies used for indoor applications is ultra wideband (UWB). UWB provides good performance within the boundaries of a small area, generally within 10-15 cm accuracy.

This paper describes on-going research into the development of a new method of autonomous navigation as applied to wireless mobile navigation. This will provide wireless autonomous mobile navigational functions to a robot on construction sites. The main goal of this research project is to integrate mobile robot's inertial navigation control unit into a UWB indoor positioning system, and identify and correct the source of errors using well-known statistical methods, such as the Kalman Filter and outliers analysis.

2 System Overview

The robot control unit should maximize the chance to reach its goal. The mobile robot should be able to measure progress in its relative frame in order to compare these measurements with its absolute frame. This is shown in Figure 1.

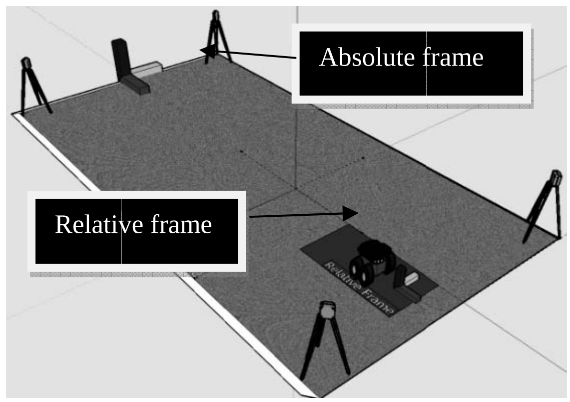


Figure 1. Absolute Frame vs. Relative Frame

Data coming from a gyroscope, an accelerometer and wheel encoders are combined to provide the robot with orientation and position in its relative frame. This is known as Dead Reckoning, or propioceptive navigation. The position accuracy is affected by slippage, uneven surface, wheel disalignment, etc. These errors can increase in severity over time, however they can be modeled, thus predicted and corrected, using statistical analysis.

Data obtained from UWB sensors and a digital compass provides the absolute position and orientation for the robot's absolute frame. This is known as exteroceptive navigation. The accuracy of these devices is limited by the resolution of the devices.

2.1 Localization systems

Estimating the position of a robot in its environment requires the knowledge of the geometry configuration (x, y, θ) for the frame attached to the robot, with respect to a local coordinate frame. This is done by determining two different reference frames. The first one is a relative frame that makes use of wheel encoders and inertial sensors such as a gyroscope or an accelerometer. The second is the absolute frame, which requires the use of external sensors, such as a UWB and a digital compass.

2.1.1 Absolute Localization

The absolute frame will be relative to the space. This frame will have also a 0,0 position, which must be related to the relative frame of the robot. It has to be updated from the robot geometry information. The 2D coordinates are determined by the readings of the UWB sensors in this research. The heading will be determined by the digital compass.

UWB System

The Ubisense system is an ultra wide-band (UWB) position system used in this research which measures Time of Difference of Arrival (TDOA) and Angle of Arrival (AOA) to achieve positioning[6]. The system consists of a number of fixed sensors which receive UWB pulses from active battery-powered tags. The fixed sensors are networked over Ethernet, and the measurement data is processed on the Ubisense software platform to give the real time track of the mobile tag [8].

Digital Compass

This sensor provides the absolute orientation in the navigation control. The main disadvantage of the digital compass is that the earth's magnetic field is affected by electromagnetic fields. This makes the use of this sensor hard to implement in indoor environment for absolute position.

2.1.2 Relative localization

The relative frame supplies the dimensions of the robot. It will be relative to the robot. This frame will have a (0,0) position at the beginning and an angle of 0. The heading of this frame will be determined by the gyroscope, and the translation will be determined by the encoders.

Odometry

Odometry is defined as the use of encoder measurements at the wheels to estimate the configuration of the robot state (position and orientation). To achieve successful autonomous mobile robot navigation, accurate odometry is essential. Localization, mapping and path-planning algorithms are all fundamental for robot navigation and all use odometry information [3].

Wheel Encoders

A wheel encoder is the essential sensor used in odometry. It is a device that is used to convert the angular position of a shaft to a digital code. It provides the distance in which a wheel has travel by basically measuring the relative distance.

Inertial Sensors

The purpose of inertial sensors is to calculate the relative change of a moving target between two consecutive sampling times, based on the measurement of acceleration and angular velocity from the inertial sensors [1]. In order for the inertial sensors to function properly, the gyroscope must be set up parallel to the direction of motion of the robot.

Gyroscope

The gyroscope measures an angular rate by picking up the signal generated by an electromechanical oscillating mass as it deviates from its plane of oscillation under the Coriolis force effect when submitted to a rotation about an axis perpendicular to the plane of oscillation. Gyroscope errors come from bias drift and noise. They are of particular importance for robot positioning because they can compensate the weakness of odometry.

Accelerometer

Accelerometer is used to measure the rate of acceleration. This sensor is used to measure the accelerations of the mobile robot. It enables the control system to know when the mobile robot is at rest.

Gyro + Accelerometer

Gyroscopes and accelerometers are used to measured rotation and acceleration. These sensors have the advantages to be self-contained, meaning that they do not need external references. The bias drift caused by the gyroscope can be fixed by using the accelerometer readings and the Kalman filter.

Relative Location + Absolute Localization

Once the absolute localization is obtained, it can be fused with the estimation of the robot's relative position so that it can correct its trajectory.

3. Source of errors

UWB

The measurement errors increase even when there is a clear open path between the UWB pulse transmitter and receiver. This system requires careful calibration before use. The signal levels for the installed environment must be calibrated. A measurement of the background noise level is required, so signal below that threshold can be excluded. [3].

Digital Compass

There are two problems associated with the digital compass as heading sensor. First, the body orientation changes either during locomotion or while standing on uneven terrain. This produces the pitch and roll of the compass, making its read-out unreliable. Second, the earth field at the compass level may be disturbed by other electromagnetic fields or distorted by nearby ferrous materials. These deterministic interferences can be categorized in two types. First, hard iron effects are caused by magnetized objects, which are at a fixed position with respect to the compass. This relative closeness should be avoided. Second, soft iron effects are caused due to the distortion of the earth field by ferrous materials [10].

Odometry

Odometry errors fall into two categories: systematic odometry errors and non-systematic errors. Usually internal systematic factors cause a rise of systematic errors, which show a biased characteristic. In contrast, non-systematic errors are independent systematic features and have an unbiased characteristic [8].

Inertial Sensors

The stochastic errors present on inertial sensors cause the subsequent numerical integrations of the measurements to exhibit an ever increasing variance. That is, when a gyro or accelerometer output is numerically integrated in a dead-reckoning navigator, the variance in the resulting position and velocity outputs grow unbounded in time [9]. This degradation of measurement accuracy propagates into the navigation solution at rates dependent on the integrity of the component sensors, the algorithms employed, and the duration of the un-aided navigation [9].

4. Proposed Error Correction

UWB

The main problem with UWB is that when there is not a clear path between the tag and sensors, it creates random points that are considered to be outliers. Consequently, outlying points have to be removed by using an outlier analysis. This analysis has to be in real time.

Outlier Removal

In statistics, an outlier is an observation that is numerically distant from the rest of a data set. In our system, this is caused by an indirect path between the sensors and tag located in the wireless robot system. There is not a mathematical definition that determines what constitute an outlier, rather it is a subjective exercise due to the variation of different samples. Many methods are used to determine whether or not an observation is an outlier. These methods are based on the mean and standard deviation of the sample. The method used as a foundation for our outlier removal is based on Grubb's test for outliers and Rosner's Test for Outliers.

Kalman Filter

The second method used to correct UWB reading is Kalman Filter. The Kalman Filter was introduced in the early 1960's and since then it has found widespread use. The purpose of the discrete–time Kalman Filter is to provide the closed form recursive solution for the estimation of linear discrete-time dynamic systems. The Kalman Filter has two steps: the prediction step, wherein the next state of the system is predicted given the previous measurement; and the update step, where the current state of the system is estimated given the measurement at that time step. The further study of these equations is left to the reader.

Digital Compass

In order to compensate the compass errors, the regression analysis is used to find a feasible pattern in a defined test bed area. This area must be located away from ferrous materials and relative closeness to power closes, which generate magnetic fluxes (Beauregard, 2006).

Wheel Encoders

This research analyzes non-systematic errors, those that result from the interaction with the surface with the wheels. The University of Michigan Benchmark (UMBmark) method is employed in the testbed so that the robot is programmed to follow a pre-programmed 4x4 square path and four spots for 90 degrees turns. This has to be completed 4 times in clockwise direction and five times in counter clockwise direction [3].

Gyroscope + Accelerometer (Inertial sensors)

The problem with this system is that path deviation at constant velocity cannot be corrected. The axis of a gyroscope also tends to drift with time, giving rise to errors. Inertial sensors allow a high rate of computation of the robot configurations, but they are not sufficient because errors are accumulated. A statistical method is required to reduce these sources of errors. Here, the Kalman Filter is used to get rid of the notorious gyroscope drift with the presented. The gyro input is a voltage measure by the sensor [12].

$$\theta_{k+1} = \theta_k + \omega_k \delta$$
$$\omega_{k+1} = \omega_k$$
$$bias_{k+1} = bias_k$$

Where θ is the angle; ω is the angular velocity; δ is the sampling period; bias is the gyro bias in angular velocity; u is the gyro output. To convert state to measurement is usually the easiest part,

$$y_k^{gyro} = \omega_k + bias_k$$

It is important to apply these steps in exact order to correct the gyroscope readings.

5 System Model

5.1 Principle

The robot has sufficient information and knowledge concerning its environment. The idea is to generate a local path using the known information, then attempt to locally, in relation to the robot, avoid any obstacles detected by the onboard sensors and return to the orientation relative to its frame.

Dead Reckoning (Relative frame)

The displacement estimates can be in the form of changes in Cartesian coordinates (i.e. x and y coordinates) or, more typically, in heading and speed or distance. With sufficiently frequent absolute position updates, Dead reckoning's linearly growing position errors can be contained within pre-defined bounds [10].

5.2 Path Planning (Absolute Frame)

The path-planning problem is usually defined as follows, "Given a robot and a description of an environment, plan a path between two specific locations. The path must be collision-free (feasible) and satisfy certain optimization criteria." In other words, path planning is generating a collision-free path in an environment with obstacles and optimizing it with respect to some criterion [11]. The robot has a short sensing range compared to the size of the region of interest. It radially senses from its position. Obstacles can block the sensing in some directions [11]. This study assumes that the robot knows its coordinates and orientation via UWB and the digital compass.

5.3 Position Algorithm (Relative + Absolute)

Once the first point is obtained in the path, this data is converted into an angle and distance relative to the robot frame. Whenever the robot reaches the target point provided by the path planning, its absolute position is calculated by using UWB readings and then passed through Kalman estimation. Then the position algorithm compares relative values of the robot, such as distance traveled and angle provided with the gyroscope, with the pre-defined bounds estimated from the UWB readings. If the robot's relative position trajectory is not within the pre-defined bounds provided by the absolute location system, then the system has to reset the robot position to the center of the bound.

6 Mobile Robot Navigation Control Architecture

Figure 2 shows the mobile robot's navigation control architecture designed in this research.

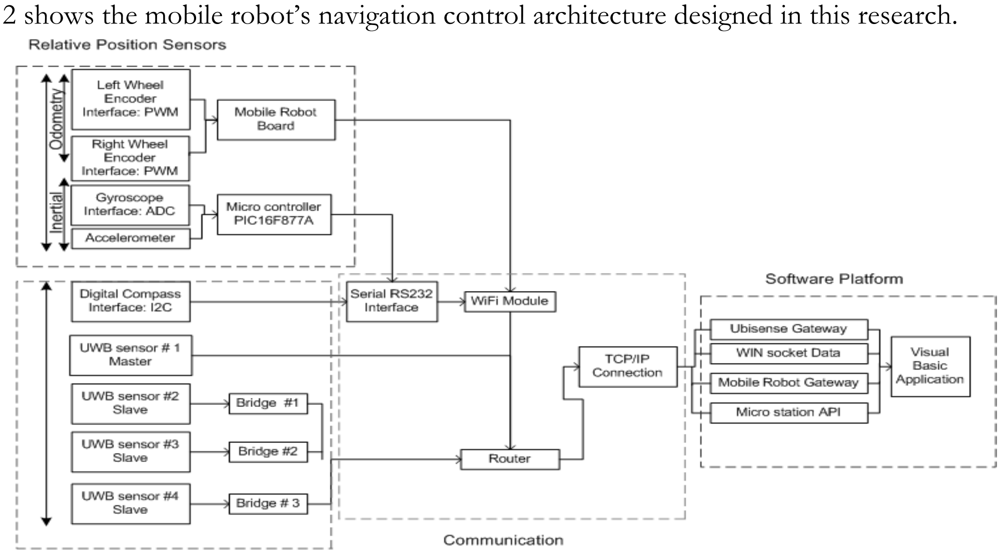


Figure 2. Mobile Robot's Navigation System Architecture

UWB

This system uses a network of sensors installed at known positions and a set of tags located in the mobile robot. The communication between different elements has two options. The first option is using an Ethernet network that basically connects all different sensors using a Cat5 Ethernet cable to a router. The second option is using a wireless interface. This option uses a set of wireless bridges connected to each sensor and respectively to the main router.

Micro controller board

The board used in this research is one of shelf solutions from Olimex™ (Figure 3). This board will be connected with the WiFi Latronix module on the mobile robot platform later. This micro controller offers a low-cost effective platform to interface with the sensor by using different interfaces. The first interface uses I2C interface that connects with the digital compass. The second interface is an Analog to Digital Converter (ADC).

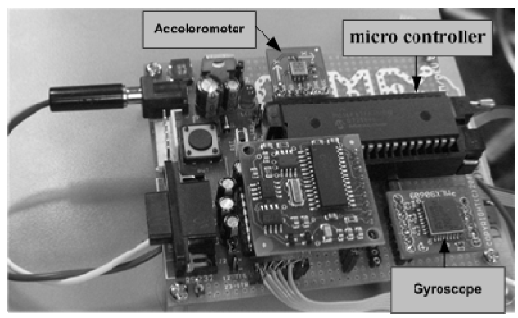


Figure 3. Micro controller board for inertial sensors

Inertial Sensors

The analog signals coming from the accelerometer and gyroscope are fed into the ADC input port from the PIC16F877A to be decoded.

Digital Compass

The compass used in this robot control is the CMPS03 built by Devantech™. It uses two methods of operation, which can be easily integrated with our microcontroller (PIC16F877A). The first option is a PWM signal that outputs a square wave. The second option is an integrated circuit interface, or I2C. This method is faster than the previous one, and allows a faster integration with the micro controller.

WiFi Module

This is a compact embedded solution that connects a UART port into its two inputs and the data can be easily accessed and controlled over a network. Mobile robot's control module uses Port A for its communication. Port B is connected to the Micro controller board.

Mobile Robot

The robot includes a series of sensors on board such as infrared, human detection, microphone, and camera among others. The one mainly used for this project is the wheel encoders that provides the data for the dead reckoning algorithm and regression analysis solution. Figure 4 summarizes the framework for this study described previously.

7. Experiments

The UWB system requires calibration within the absolute and relative frames such that a common point of origin (0,0) is established. Using a Total Station an absolute reference point is established. Each sensor has to be calibrated independently to this point of origin.

Test # 1 UWB- Average + Kalman: A buffer array, which has 5 slots for each data X and Y, is created. Data enters the system every 0.2 seconds. The data is then put into the 5 slot array, an average is calculated, and the array shifts to accommodate incoming data. Each calculated average is stored as a temporary variable and fed into the Kalman filter. The results are shown in Figure 5.

Test # 2 UWB-MidValue + Kalman : This method is similar to the previous, however, the median value is chosen rather than the mean. The result can be observed in Figure 6.

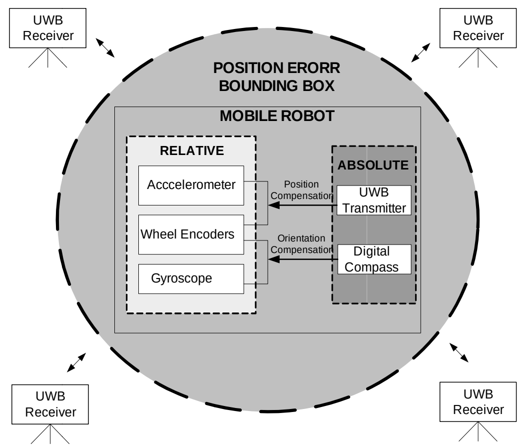


Figure 4. Framework for error compensation between relative and absolute sensors

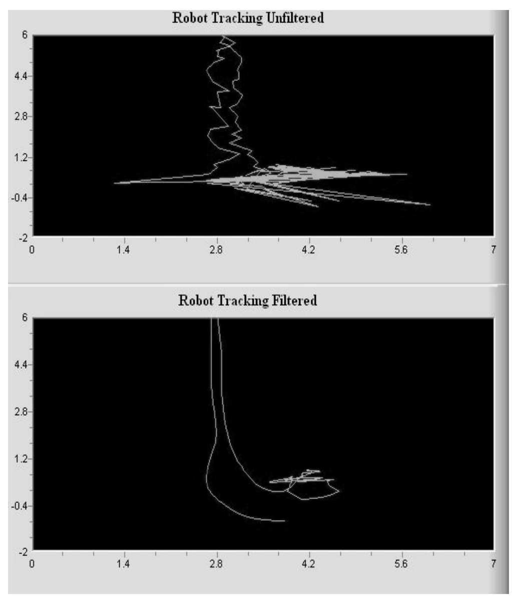


Figure 5. UWB- Average + Kalman

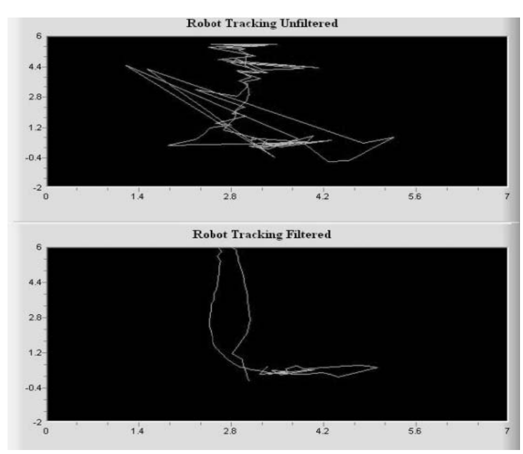


Figure 6. Test # 2 UWB-MidValue + Kalman

Test # 3 UWB- Outliers Removal + Kalman: While similar to the previous methods, this test includes the calculation of a threshold where outliers are removed based on Outlier analysis described in Ronsner's method(Figure 7).

Three methods tested show a significant improvement in UWB localization, and Test # 3 provided the best results.

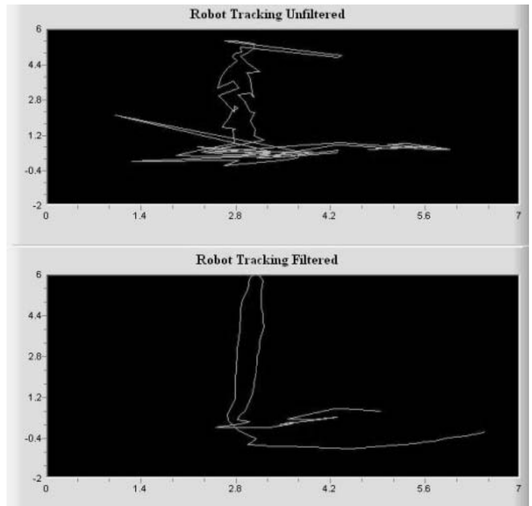


Figure 7. UWB- Outliers Removal + Kalman

8 Conclusions

This paper introduced an on-going research project developing indoor robot wireless navigation leading to the implementation of a suitable localization system within indoor construction environments. A framework has been introduced to improve the mobile robot's navigation in construction processes. The proposed framework introduces methods to detect and correct errors from various sensors including both internal and external sensors. Each sensor's error attributes are identified through intensive and extensive lab tests. Initial results from this research have shown an improvement in real time for the UWB localization system. This improvement was achieved by combining different statistical analysis. Three different experiments were conducted with results supporting the above conclusions.

Further investigation is being conducted to address additional sensor related errors based on the framework which has been established in this study. It is expected to lead to the improvement of the remaining sensors integrated within the system. These improvements will lead a better, more suitable localization system that can be applied to various indoor construction automation applications.

References

[1] Tao. Y, Hu. H, Zhou. H. (2005)."Integration of Vision and Inertial Sensors for Home-based Rehabilitation". IEEE International Conference on Robotics and Automation, Barcelona, Spain InerVis the 2nd Workshop on Integration of Vision and Inertial Sensors,.

[2] Barshan, B. and Durrant-Whyte, H.F. (1993). "An Inertial Navigation System for a Mobile Robot." Proceedings of the 1993 IEEE/RSJ International Conference on Intelligent Robotics and Systems, Yokohama, Japan, July 26-30, pp. 2243-2248.

[3] Borenstein, J. and Feng, L. (1994). "UMBmark & A Method for Measuring, Comparing, and Correcting Dead-reckoning Errors in Mobile Robots." Technical Report, The University of Michigan UM-MEAM-94-22, December.

[4] Bock, T and Kreupl, K. (2004). "Procedure for the implementation of Autonomous Mobile robots on the Construction site," Proceedings. 21st International Symposium on Automation and Robotics in Construction, Jeju, Korea, 304-309.

[5] Williams, C., Cho, Y., Youn, J., Jung, U. (2007) Wireless Sensor-driven Intelligent Navigation Method for Mobile Robot Applications in Construction, ASCE, 2007 International Workshop on "Computing in Civil Engineering, ASCE, Pittsburgh, PA, ASCE, 640-64. pages
[6] Ringstad, G and Enge.P. (2008)"GPS and UWB for indoor navigation". Stanford University. <http://waas.stanford.edu/~wwu/papers/gps/PDF/guttormion01.pdf.>
[7] Martinelli, A. and Siegwart, R. (2003). "Estimating the odometry error of a mobile robot during navigation. In European Conference on Mobile Robots," Warsaw, Poland, Sept. 4-6.
[8] Moore, T. Hill, C., Hide C., Ochieng W.Y., Feng S., Aguado E., Ioannides, R., Cross P. and Lau L. (2007). "End-to-End Testing of an Integrated Centimetric Positioning Test-Bed," Proceedings of ION GNSS 20th International Technical Meeting of the Satellite Division, Sept. 25-28.Fort Worth, TX, pp.1128-1138.
[9] Wall. J. and Bevly, Y .(2008). "Characterization of Inertial Sensor Measurements for Navigation Performance Analysis". <gavlab.auburn.edu/paper.php?id=22> (Dec. 12, 2008).
[10] Beauregard, S.(2006). "A Helmet-Mounted Pedestrian Dead Reckoning System", 4th International Forum on Applied Wearable Computing.
[11] Sedighi, K., Ashenayi, K. Manikas, T.W. Wainwright, R., Heng-Ming Tai . "Autonomous local path planning for a mobile robot using a genetic algorithm." Evolutionary Computation, CEC2004.Volume: 2, pages: 1338- 1345.
[12] T. Tong. "Kalman Filter Made Easy". http://openuav.astroplanes.com/library/ docs/writeup.pdf> (Dec. 10, 2008).

Dynamic Simulation for Assessment of Impact of Highway Lane Closure on Traffic

Seokyon Hwang

Assistant Professor, Department of Civil Engineering, Lamar University, P.O. Box 10024, Beaumont, TX 77710. PH (409) 880-7557; FAX (409) 880-8121; e-mail: seokyon.hwang@lamar.edu

Abstract

Improvement of existing highway systems, such as maintenance, rehabilitation, and reconstruction, is accompanied with lane closure at work zones. Such lane closure causes road user costs for traffic congestion and delays at the work zones. A robust method is developed in this research to support strategic decision-making on lane closures. At this stage, the research focuses on a common situation—closing a segment of one lane in one direction on a four-lane highway. The method divides a work zone into arriving-, entering-, passing-, and leaving-zone and classifies the flow of vehicles as arriving-, in-, through-, and out-flow. Thereby, it accounts for traffic congestion and delay due to speed reduction, upstream queue, and speed recovery. Based on dynamic simulation technique, the method simulates the flow of vehicles and measures congestion and delay at work zones. The new method would benefit engineers in planning highway system improvement projects.

Keywords: Highway, Reconstruction, Lane closure, Congestion, Delay, Simulation

Introduction

In the United States, public expenditures have increased in recent years to renew and revitalize aging highway systems (U.S. DOT 2006) where most existing systems are deteriorating at a fast pace (ASCE 2002). Federal Highway Administration (FHWA) and state departments of transportation (DOTs) estimated the need of a huge amount of average annual capital investment in highways and bridges for the 20-year period 2005–2024: $78.8 billion for maintenance of the systems and $131.7 billion of maximum economic investment at all government levels (U.S. DOT 2006).

In addition to direct capital investment, highway improvement projects have a broader impact on the society. Such projects are inevitably accompanied with lane closure at work zones, either partially or completely. For instance, a certain length of stretch of a road needs to be closed to repair or replace deteriorated concrete pavement, which reduces the capacity of the road. As a result, traffic demand on the road can exceed the reduced capacity, which can incur traffic congestion and delay caused by speed reduction and queue of vehicles at the work zone (Carr 2000), and even complete failure of highway operations in the peak time periods (Martinelli and Xu 1996). This phenomenon ultimately can incur a considerable amount of public cost, such as delay cost and loss of business, and increased safety risk.

A robust method was developed in this research based on dynamic simulation technique to quantify the impact of lane closures on traffic at work zone. It evaluates the impact by means of measuring travel time of vehicles. Dividing a work zone into four sub-zones (arriving-, entering-, passing-, and leaving-zone), the method accounts for the effect of speed reduction, upstream queue, and speed recovery on traffic flow. The method simulates the flow of vehicles (arriving-, in-, through-, and out-flow) and measures travel time of vehicles passing through a work zone. Thereby, it accounts for feedback on dynamics among the flows.

The study presently focuses on one of the most common situations—closing a segment of one lane in one direction (single-lane closure) on a four-lane divided highway. Discussion in this paper focuses on the following findings: the strategies for the development of the method, a preliminary dynamic simulation model, and an application example for demonstration of the method.

Lane Closure Decision and Traffic

Theoretically, an unlimited number of ways of lane closure are possible to configure a work zone in a highway improvement project. Given a four-lane divided highway, the following exemplifies a few possible ways to configure a work zone: (1) closing one lane in one direction (single-lane closure), (2) completely closing two lanes in one direction and diverting all vehicles in that direction to the roadway in the opposite direction, and (3) completely closing two lanes in one direction and detouring all vehicles in that direction. Too many options, on the other hand, create the challenge of evaluating each option and selecting the best performing closure strategy. This pronounces the need for effective methods for quantification of impact of each strategy to evaluate various possible options for lane closure.

In the study of work zone problem, previous efforts have been concerned with a variety of issues associated with planning and operation of a work zone. Some representative issues include the following: the analysis of factors causing traffic congestion and delays at a work zone, the calculation of delays and user costs, the optimization of traffic control design and practices, the design and development of optimal control devices, and the optimal configuration of a lane closure strategy. Attempting to improve decision-making process while dealing with such challenging issues, various methods and approaches have been developed. The following briefly discusses a few representative studies.

Much effort has been devoted to improving accurate estimation of road capacity, delays, and user costs. Lane closure naturally decreases the design capacity of a road, thus, estimating capacity loss is important when planning a lane closure. Krammes and Lopez (1994) developed a computer model to estimate road capacity during maintenance work. Noting that reduced capacity at a work zone incurs delays caused by speed reduction and congestion, estimating traffic delays, user costs, and vehicle operating costs has been intensively studied (Mammoth and Dudek 1984; Dudek et al. 1986; Krammes et al. 1987; Carr 2000). In an attempt to optimize work zone length, traffic control design, and practices, various methods have been developed to compare and evaluate different lane closure strategies (Mahmassani and Jayakrishnan 1988; Chien et al. 2002; Chen and Schonfeld 2005).

Meanwhile, unlike the above technical solutions, administrative solutions also have been sought for. In an effort to mitigate traffic congestion and resulting impacts, the government agencies have adopted a variety of contracting methods, such as A+B contract and A+B+I/D (Herbsman 1995). These methods incorporate road user costs in project costs, aiming to reduce disruption to traffic so as to minimize the overall societal costs.

Method of research

Problem Context

Depending on traffic demand, congestion tends to occur during a certain time period of a day (Dudek et al. 1986). In other words, traffic congestion may occur in different scales under various situations at a work zone. For instance, demand may increase during the commuting time periods in urban freeways. Thus, it is important to estimate traffic congestion given various volumes of traffic over time. Measuring the number of vehicles traveling through a work zone for a particular time period, it is possible to evaluate the effect of lane closure on the road capacity at the work zone and the operation of the road during the time period.

With regard to traffic speed at a work zone, there are a few prevailing phenomena. (1) If there is congestion at a work zone, the speeds of vehicles decrease as vehicles approach to the area. (2) Even if there is no congestion, vehicles may have to slow down to comply with the limit of speed at the work zone imposed by law for safety of road workers and drivers. (3) The speeds of vehicles in a work zone eventually depend on how fast vehicles move out of the area and accelerate to normal speed to leave the work zone area. (4) The pattern of flow of a vehicle at and near a work zone affects the speed of the vehicle. (5) The length of a work zone can become an influential factor affecting congestion and delays (Martinelli and Xu 1996; Chen and Schonfeld 2005); thus, the length of a work zone should be taken into consideration when estimating the speed of a vehicle traveling the work zone and the amount of time the vehicle consumes in the area.

Behaviors of Vehicles and Work Zone Decomposition

Concerning the speed of traffic at a work zone, traffic normally exhibits three patterns of speed—slows down before the work zone, maintains reduced speed or even further slows down in the work zone, and slowly speeds up after the work zone. Considering these patterns, this research divided a work zone and its vicinity into arriving-, entering-, passing-, and leaving-zone. Meanwhile, the actual capacity of a work zone is determined by the flows of vehicles at four sub-zones discussed earlier. For instance, congestion is caused at entering-zone by the reduced flow of vehicles into the passing-zone. Therefore, the flows of vehicles, which are defined as arriving-, in-, through-, and out-flow, are basic information required to explain congestion and delays of traffic at a work zone that are caused by speed reduction, upstream queues, and speed recovery.

The following briefly explains each of the four flows. The arriving-flow represents the normal flow occurring before a work zone. Both in-flow and through-flow are influenced by congestion caused by accumulating vehicles in the passing-zone, which is due to reduced capacity and mandatory slow driving for safety. Finally, out-flow is dependent on congestion occurring at leaving-zone in which vehicles are recovering normal speed right after the work zone.

System Dynamics Modeling

System dynamics modeling is known to be a useful technique for translating the real-world systems and abstracting those into model elements by which the dynamics in the systems can be effectively explained. Elements and mechanisms of a model as well as flow of information among elements can be explicitly represented in a model (Hannon and Ruth 1997). The properties allow overcoming a significant shortcoming of the existing approaches to the lane closure problem. Existing methods generally deal with the delays due to congestion and the delays due to speed reduction in a work zone separately. System dynamics modeling allows modeling the dynamics existing in the system in a model, so that the resulting model can simultaneously deal with congestion, speed reduction, and queues of vehicles.

A system dynamics model comprises three kinds of system variables (Hannon and Ruth 1997). State variables indicate the current status of the system and represent the states of stocks flowing in the system. Flow variables represent changes to the states of stocks that happen as simulation runs. Transforming variables influence flow variables so as to make changes to the state variables.

Dynamic Simulation ModeL

System Variables and System Dynamics Model

The flow of traffic in and near a work zone is very dynamic. The dynamics can be attributed to feedback effect caused by the flows of vehicles at four sub-zones that are affected by congestion, speed reduction, and queues of vehicles. Table 1 presents the variables of the present model: state variables (number of vehicles at each zone), flow variables (flow of vehicles through each zone), and transforming variables (factors and controls influencing flows of vehicles). The demand of vehicle per minute (vpm) represents the number of vehicles arriving at a work zone per minute. Its distribution, both normal and reduced, can be obtained from vehicles per hour (vph) traveling the road. The distribution normally changes over time depending on actual demands at different time periods. In order to account for random arrivals of vehicles, the randomness is introduced to arrivingFlow and the random effect influences other flows followed by; more details are discussed in the following section. The maximum number of vehicles (maxVeh) represents the largest number of vehicles that are allowed to be in passing-zone any time periods. This parameter is assumed to be estimated by engineers in proportion to the length of the work zone (L), considering safety requirements such as limited work zone speed, road performance conditions, work layout, etc.

Figure 1 illustrates the developed system dynamics model, Dynamic-T (System Dynamics AssessMent of Impact of lane Closure on Traffic). One part of the model (double-lined box) simulates vehicle flow and the other part (single-lined box) collects cumulative number of vehicles that have passed through each flow. Using the flow rates of vehicles between the sub-zones, the model simulates and analyzes the flow of vehicles across the work zone. It also measures the volume of traffic passing through the work zone over time. A graphical dynamic modeling and simulation tool, StellaTM (iseesystems Inc. 1985), was used to construct and simulate the model.

Table 1. Variables of the Simulation Model

Type of variable	Variables
State variable	*vehArrived, vehInWZ;* *TOTALvehEntered, TOTALvehArrived, TOTALvehLeft*
Flow variable	*arrivingFlow, inFlow, outFlow;* *arriving, entering, leaving*
Transforming variable	*vpm, L, maxVeh, speedReduction, OnOff, AC1, AC2*

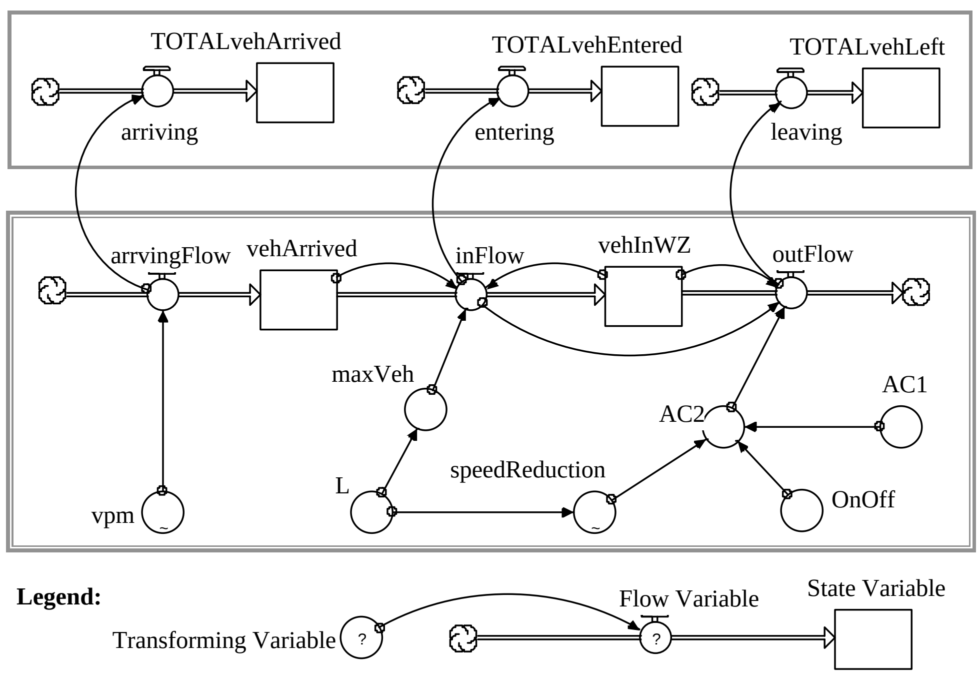


Figure 1. System Dynamics Simulation Model (Dynamic-T)

Simulation Algorithm

Arrival rate of vehicles normally is known to be random and represented by a certain distribution. While considering randomness, many queuing simulations adopt the Poisson distribution for arrival rate, assuming the average rate per interval time is known and fixed. Known the average rate from historical data normally is reported on a one-hour interval basis. The rate changes in time depending on various levels of demand at different time periods. In the new method, normal distribution is used to represent the arrival rate. It is not effective to apply a fixed arrival rate considering continuous change in demand, where the primary interest of the study is to assess continuously the impact of lane closure during a day, i.e., twenty four hours. Figure 2 presents the algorithm of the present simulation model. The flow of vehicles at each zone, except for *arrivingFlow*, is controlled by the state variables at *vehArrived* and *vehInWZ*, as well as the transforming variables, related to the flows. The *arrivingFlow*, the number of vehicles arriving per minute, is determined by the demand of vehicle per minute, *vpm*. Unless there is a natural decrease in demand or a planned decrease arranged by diverting vehicles before a work zone, *arrivingFlow* is clearly independent of any variables, but *vpm*. *vpm* data is normally acquired from historical data such as highway capacity report, which takes into consideration the ratio of trucks out of total demand.

The *inFlow* can be delayed by the state variable. *vehInWZ*, because the simulation does not allow more than a certain number of vehicles per unit length, *maxVeh*, in *vehInWZ*. For instance, engineers based on historical data may estimate that the number of vehicles in a work zone should not exceed eighty vehicles per mile for safety. The *inFlow* is affected by *arrivingFlow*, because it determines the state of *vehArrived* that controls the flow rate of vehicles at *inFlow*. The *inFlow* is also influenced by *outFlow*, because *outFlow* determines the state of *vehInWZ*. The through-flow, one of the four flows defined earlier, is not implemented in the simulation, because it is a conceptual flow occurring in *vehInWZ*, which is governed by *outFlow*. The more vehicles are in the *vehInWZ*, the longer a queue develops in entering-zone, and the less number of vehicles are allowed to flow at *inFlow*.

```
vpm = GRAPH(time)

vehArrived(t) = vehArrived(t - dt) + (arrvingFlow - inFlow) * dt
vehInWZ(t) = vehInWZ(t - dt) + (inFlow - outFlow) * dt
arrvingFlow = normal (vpm, vpm*.05)
inFlow = IF (vehInWZ<=maxVeh) THEN (IF (vehArrived>(maxVeh-vehInWZ)) THEN (maxVeh-vehInWZ)
         ELSE (vehArrived)) ELSE (0)
outFlow = IF(inFlow<=AC2) THEN inFlow ELSE IF vehInWZ>AC2 THEN AC2 ELSE vehInWZ

TOTALvehArrived(t) = TOTALvehArrived(t - dt) + (arriving) * dt
TOTALvehEntered(t) = TOTALvehEntered(t - dt) + (entering) * dt
TOTALvehLeft(t) = TOTALvehLeft(t - dt) + (leaving) * dt
arriving = arrvingFlow
entering = inFlow
leaving = outFlow

AC2 = IF OnOff=1 THEN AC1*speedReduction ELSE AC1
speedReduction = GRAPH(L)
maxVeh = 80*L
```

Figure 2. Simulation Algorithm

It is assumed that the road capacity at a work zone can be reduced due to the limit of speed the zone, *speedReduction*, and the decreased vehicle speed for the loss of one lane, *AC1*. Both parameters are to be determined by engineers based on existing historical data, for instance, published data by DOTs. The *OnOff* allows switching s*peedReduction* from active to inactive, and vice versa, in calculating the augmented capacity reduction, *AC2*. The *AC2* is determined by considering speed limit, capacity reduction for reduced number of lanes, and the length of lane closure.

Output and Analysis

The simulations of a work zone with the model produce many informative outputs, including the number of vehicles passing each sub-zone during individual periods and over time. Interpreting these data, engineers can produce important information with regard to the impact of lane closure on traffic and highway system operation during construction. Both are useful to estimate service level of the road for different demand over time by measuring traffic congestion and delays at different time periods of a day. In this manner, engineers can evaluate each lane closure strategy among many possible candidates by quantifying its impact on traffic. To name a few decisions that can be effectively made by using such information, it includes the length of lane closure, time to impose speed limit, if flexibly operated, and time to start work at a work zone. Additionally, the data can provide information on the volume of vehicles to be detoured to maintain a certain level of service.

A Hypothetical Example

A hypothetical situation was created for demonstration of the new method. A set of input data—road capacity, speed of vehicle, and traffic demand—was prepared based on published data in the Highway Capacity Manual (TRB 1985). Other parameters were assumed to be estimated by engineers.

Situation

A segment of a lane needs to be closed to replace existing pavement on a four-lane divided highway. At the work zone, traffic in two lanes merges into one open lane, and diverting or detouring traffic is not considered. Recent data samples show that the average hourly demand varies over time during a day. Vehicles using the road comprise 25% of truck and 75% of cars. Figure 3 depicts the relationship between the length of lane closure (L) and the reduction of vehicle speed (speedReduction), for which data is acquired from a recent survey. Figure 4 (a) presents the distribution of traffic demand during a day. The demand of vehicle per minute is calculated from vehicle per hour. According to historical hourly capacity data, the average capacity of one lane in one direction is 1500 vph that is equivalent to 25 vpm, without any speed limitation in the work zone. Engineers consider two lane closure strategies in terms of the length of closure—2 miles and 5 miles. Running the simulation for 24 hours (1,440 minutes) with 1 minute interval, the following results were obtained.

Results

Figure 4 (b) and (c) illustrate the results of two cases—2 mile and 5 mile long lane closures. Briefly discussing the results, the longer lane closure resulted in more congestion and delays. For both cases, traffic started building up a queue from the moment that demand exceeds capacity. The queue was transferred to the next time periods and started to decrease when demand becomes smaller than capacity. With 2 mile long closure, traffic congestion was observed two times, 7 - 10 a.m. and 6 - 11 p.m., whereas congestion was observed almost a whole day with 5 mile long closure, 7 a.m. - 12 a.m. Decision of choice can be made depending on the threshold of acceptable congestion level that is set by engineers.

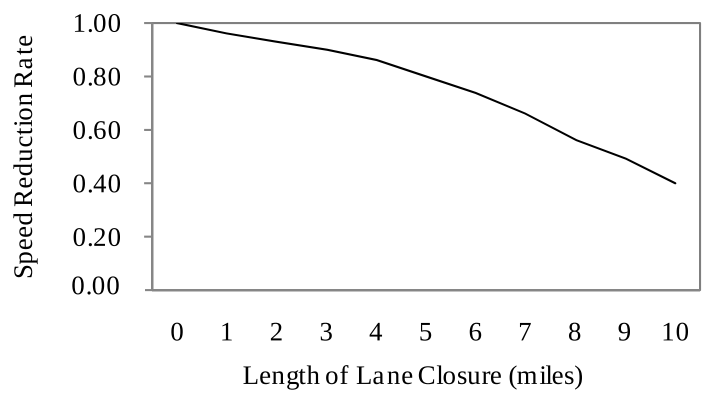
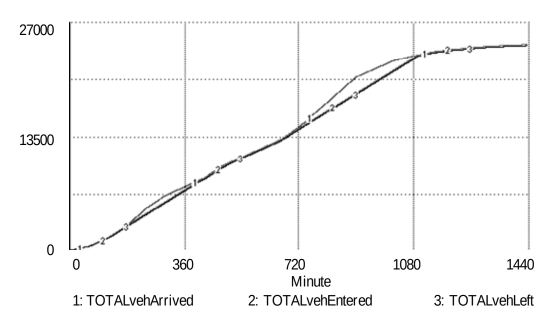


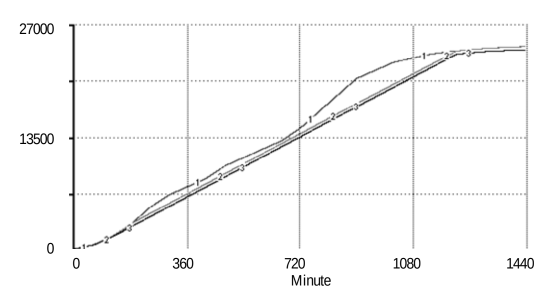
Figure 3. Relationship of *speedReduction* and Length of Lane Closure (*L*)

Minutes	Time of Day	Vehicle Per Hour (vph)	Vehicle Per Minute (vpm)
0	A.M. 5	500	8.33
60	6	1000	16.67
120	7	1500	25.00
180	8	2000	33.33
240	9	1500	25.00
300	10	1000	16.67
360	11	1000	16.67
420	P.M. 12	1500	25.00
480	1	1000	16.67
540	2	1000	16.67
600	3	1000	16.67
660	4	1500	25.00
720	5	2000	33.33
780	6	2000	33.33
840	7	2000	33.33
900	8	1000	16.67
960	9	1000	16.67
1020	10	500	8.33
1080	11	500	8.33
1140	A.M. 12	300	5.00
1200	1	200	3.33
1260	2	200	3.33
1320	3	100	1.67
1380	4	100	1.67
1440	5	300	5.00

(a) Demand Over Time



(b) *L*=2 miles



(c) *L*=5 miles

Figure 4. Results of an Application Example

Conclusions

Highway projects for keeping or enhancing the existing systems presents a broad impact to the society. Accompanied lane closure at a work zone seriously causes the capacity reduction, which in turn incurs speed reduction and resulting traffic congestion and delays. Such undesirable situation creates tremendous amount of road user costs, in addition to construction costs. Thus, a strategic decision on lane closure is needed to reduce the congestion and delays. When planning improvement projects, engineers need to be able to estimate the amount of impact of various configuration of lane closure. A robust method was developed based on system dynamics simulation technique. The method is envisioned to be a useful tool for decision-making in designing and operating lane closure at a work zone. The method is effective to estimate congestion and delays by measuring and analyzing traffic flows across a work zone. It is effective in identifying and measuring traffic congestion and delays over time under various situational demand of traffic. In addition, the simulation model simplifies the process of the evaluation of lane closure strategies.

Developed method requires situational demand over time, which is manually provided for simulations. In future, the task needs to be automated so that engineers can test a variety of situational demand without paying much effort on creating data sets. The present research initiative can also be extended further to encompass more lane closure strategies and configuration. Developing a model for each of those may require additional modeling activities. However, findings from the present study can provide a strong foundation. Additionally, the problem scope can be extended to look into the work zone problem from a broader perspective. In order to find a solution to minimize the broader impact, it is necessary to take into consideration direct construction costs, traffic control costs, accident costs, road user costs, and safety and productivity on the job site.

References

[1] American Society of Civil Engineers (ASCE). (2002). "Report card: Reports and statistics." <http://www.asce.org/reportcard/>.
[2] Carr, R. I. (2000). "Construction congestion cost (CO) basic model." J. Construction Engineering and Management, 126(2), 105-113.
[3] Chen, C., and Schonfeld, P. (2005). "Work zone lengths for a four-lane road with an alternate route." J. Transportation Engineering, 131(10), 780-789.
[4] Chien, S., Tang, Y., and Schonfeld, P. (2002). "Optimizing work zones for two-lane highway maintenance projects." J. Transportation Engineering, 128(2), 145-155.
[5] Dudek, C. L., Richards, S. H., and Buffington, J.L. (1986). "Some effects of traffic control on four-lane divided highways." Transportation Res. Rec. 1086, 20-30.
[6] Hannon, B., and Ruth, M. (1997). Modeling dynamic biological systems. Springer, New York.
[7] Herbsman, Z. J. (1995). "A+B bidding method - hidden success story for highway construction. " J. Construction Engineering and Management, 121(4), 430-437.
[8] iseesystems Inc. (1985). Stella v 8.1 <http://www.iseesystems.com>.
[9] Jaraiedi, M., Plummer R. W., and Aber, M. S. (1995). "Incentive/distincentive guidelines for highway construction constracts." J. Construction Engineering and Management, 121(1), 112-120.
[10] Krammes, R. A., and Lopez, G. O. (1994). "Updated capacity values for short –term freeway work zone lane closures." Proc., TRB 73rd Annual meeting, Transportation Research Board, Washington, D.C.
[11] Krammes, R. A., Marsden, B. G., and Dudek, C. L. (1987). "Micor-computer tool for freeway work zone planning and scheduling." J. Transportation Engineering, 113(4). 348-356.
[12] Mahmassani, H. S., and Jayakrishnan, R. (1988). "Dynamic analysis of lane closure strategies." J. Transportation Engineering, 114 (4), 476-496.
[13] Mammoth, J. L., and Dudek, C. L. (1984). "Queue and cost evaluation of work zones (QUEWZ)." Transportation Research Board, Washington, D.C., 12-19.
[14] Martinelli, D. R., and Xu, D. (1996). "Delay estimation and optimal length for four-lane divided freeway workzones." J. Transportation Engineering, 122(2), 114-122.
[15] Transportation Research Board. (TRB). (1985). Highway capacity manual, Special Rep. No. 209, TRB, Washington, D.C.

[16] U.S. Department of Transportation (U.S. DOT). (2006). Status of the nation's highways, bridges, and transit: 2006 Conditions and performance, <http://www.fhwa.dot.gov/policy/2006/>.

Assessment of Performance Metrics for Use of WSNs in Buildings

Won-Suk Jang[1] and William M. Healy[2]

[1]Building and Fire Research Laboratory, National Institute of Standards and Technology, Gaithersburg, MD 20899-8632; PH: (301) 975-5855; FAX: (301) 975-5433; e-mail: wonsuk.jang@nist.gov
[2]Building and Fire Research Laboratory, National Institute of Standards and Technology, Gaithersburg, MD 20899-8632; PH: (301) 975-4922; FAX: (301) 975-5433; e-mail: william.healy@nist.gov

Abstract

This paper is part of an effort to develop measurement systems and methods to better predict the performance of wireless sensor and control networks in building applications. Because of the variability in application requirements, the development of measurement methods to assess the performance of wireless sensor networks in buildings is extremely challenging. This paper presents the key challenges in using wireless technology in practical building applications through the results of a literature survey and interactions with building industry professionals. It is anticipated that the findings will provide potential users of wireless technology a clear metric as to how the technology will perform in a particular, and to understand the barriers to adoption of wireless sensors in buildings when appropriate.

Introduction

Wireless technology facilitates easy deployment of sensors throughout various applications in building automation, construction, and structural health monitoring. However, many engineers and operators who may consider using those sensor systems do not have a clear metric to describe how wireless systems will work in particular applications. These measurement needs are stated in an assessment of the United States' measurement system as "Potential end-users of wireless sensor networks have shown reluctance towards using them in a wider range of applications because of uncertainty in the reliability of the wireless links [Swyt, 2007]." Furthermore, a recent DOE roadmap entitled "Advanced Sensors and Controls for Building Applications" recommended the development of an Operational Test and Evaluation Program for sensor systems that would "provide systematic and comparable testing, employ and develop testing protocols that are standards based, and help to identify environmental vulnerabilities and operational limitations.[Brambley et al., 2005]" Owing to the variability in both building types and application requirements, the development of measurement methods that can be used to assess the performance of wireless sensor networks in buildings is extremely challenging. As the first part of the effort, this paper presents the key challenges in using wireless technology in building applications to ensure that the work addresses the critical barriers to the use of wireless sensors and controls when they can provide a benefit to the building owners and occupants.

Background

To identify these needs, a research team at the National Institute of Standards and Technology (NIST) surveyed the literature and interviewed people involved in building operations and wireless products. Interviews were performed of people who represent large building automation companies, wireless equipment manufacturers, engineering consulting and design-build firms, building maintenance departments, and research organizations. These people were typically asked about 1) the challenges or concerns that they see in using wireless sensors and controls in buildings, and 2) the types of measurements that they could envision that would give users more confidence in the performance of a wireless system.

This paper addresses these obstacles to the use of wireless technology in buildings by developing measurement methods and testbeds that will enable users to predict the performance of a wireless system for a building application. While each application is different, it is anticipated that a consistent set of metrics will provide useful information in designing the layout of the wireless network for a desired level of service.

Practical Challenges

This section highlights the most important, practical issues that can address a useful metric for designs and applications of wireless sensor in a building.

1) Cost

A major advantage in using wireless sensors and controls over wired systems is the decreased installation and maintenance cost. In wired systems in a large building, cost issues become prominent when the number of wires and complexity of the network increases. The installation cost for wired systems was reported as $7.21 and $2.20 per linear meter of wiring in existing construction and new construction, respectively [Kintner-Meyer and Brambley, 2002]. In addition, it has been reported that the typical wiring process accounts for 75% of the installation cost of sensor networks in a structural health monitoring system [Lynch, 2007]. While the installation cost depends upon the type of construction, the size of the building, and the need for radio repeaters, promised declines in the prices of the radio hardware will make wireless more competitive.

The number of nodes and repeaters is the key factor in assessing the cost of a wireless system. The number of nodes will be based on the desired resolution in a particular application, while the building construction, size, and wireless transmission scheme will dictate the need for repeaters. Designers may need to make tradeoffs to achieve a desired price point for a particular installation.

Overall, it is obvious that an installer of a wireless system must see how much this system will cost over a wired system that performs the same function. A clear metric on assessing the costs of a wired system versus a wireless system would greatly help users see financial benefits or drawbacks of a wireless system.

2) Reliability

Reliability of the wireless system is one of the major concerns of potential users of wireless technology in buildings. At its most basic level, the same level of reliability with a wireless system is expected as is seen with wired systems, while its reliability requirement may either be more or less severe in a broader range of applications.

Defining reliability is itself a difficult endeavor. The following discussion will describe different aspects of reliability and factors that affect reliability.

Accuracy

Measurement accuracy is not specific to wireless sensor networks but is a concern with all sensor devices. Traditionally, internal or external noise imposes inherent limitations on the accuracy of sensors and equipment. The noise may affect sensor readings by modifying the analog signal generated by the transducer. Typically, sensor readings will be converted from analog to digital format at the sensor node before being transmitted wirelessly. Because the data are transmitted digitally, the change in accuracy of the measured value because of noise interfering with the data transmission is less of a concern. In addition, data corruption or distortion caused by environmental impact will typically result in values that are noticeably in error.

One positive characteristic of wireless sensor networks with respect to accuracy is node redundancy. High node density may compensate for overall uncertainty caused by the noise and non-ideal conditions, and can increase the overall accuracy of the measurement in a particular areas. For instance, a cluster-based topology can be employed to collect sensory data at the same local area. While this topology presents the problem of knowing which sensor's data are closest to being correct, it does present the possibility of obtaining data when certain sensors become defective or in using statistical methods to determine the best estimate of the true value of a parameter. Thus, densely distributed wireless networks could help increase the overall accuracy by allowing redundant measurements.

Signal Coverage through Building Materials

Different construction types and the resultant coverage of the wireless signal are significant factors for the reliability of a wireless link. In most cases, the maximum allowable distance between the transmitter and the receiver is specified when there are no obstructions between the two radios (e.g., in an open field). However, the actual prediction of signal propagation is much more complicated because walls, floors, ceilings, and furnishings are often present in the indoor environment. Different construction materials also attenuate the signal to varying degrees, and even tend to stop propagation of radio-frequency (RF) waves. Further complicating matters is the fact that different frequencies and transmission power levels may respond in different manners in different buildings. Without clear understanding on the allowable ranging distance in real applications, users may both place more repeaters than necessary (and, hence, increase the cost of their system) or run the risk of locating sensors in places where their data will not be reached.

Interference

A low-power wireless device is potentially vulnerable to interference from other wireless technologies that have much higher power within the same industrial, scientific, and medical (ISM) band. Typically, many of the devices being constructed as part of wireless sensor networks operate in the 2.4 GHz band because of its worldwide usage. While the risk of interference is minimized by various techniques of signal communication, there still exists the potential for interference from other equipment considering the large number of devices emitting at similar frequencies. From interviews, end-users expressed concerns that the devices will not report healthy data at certain times because of the possibility of interference. A standard measurement technique of wireless system performance given a typical interference pattern could help users understand the effects of other devices on their wireless systems.

Latency

Acquiring data in real-time is considered an application-specific issue in wireless sensor network. Inherent features of WSNs, such as dynamic topology, lossy links, limited bandwidth, and channel variations, often limit the real-time performance. Different demands on end-to-end latency are required because most applications have different expectations that data acquired from a wireless sensor should be made available to a receiver within a reasonable period of time. While the end-to-end latency is most crucial for certain applications, e.g. emergency response, quality-of-service (QOS) is usually a higher priority for many applications. Thus, most solutions for the real-time issue involve trade-offs between power management and latency, and careful management of transmission control and topology should be provided.

Fault Tolerance

In randomly deployed wireless sensor nodes, the failure of individual components, such as a node, network, or sink, is unavoidable. In the design of sensor networks for building application, it is important to identify the failures that will affect the overall performance of the system and that will degrade the confidence level in the measurements. In building applications, a fault can be classified by three categories: 1) node faults, 2) network faults, and 3) sink faults. First, various sources of node faults can be identified under harsh environmental conditions. A variety of extreme conditions can cause antenna failures, circuit failures, and battery leakage, which will lead to poor performance of WSNs. Second, a network fault is another common fault in a WSN since the high density of deployed nodes will increase the chance of individual communication failures. Third, sinks, the points where data are collected, are subject to faults. Typically, the power supply and network infrastructure are the main components for sink faults that will cause the malfunction of the entire sensor network.

There are a number of metrics that one can use to assess or predict reliability:

Received Signal Strength Indication (RSSI)

Received Signal Strength Indication (RSSI) is a term used to describe a measurement of the power present in a received radio signal. RSSI is calculated at the radio chip on the receiver and provides useful implication of network link quality. The drawback of RSSI as an indicator of the reliability of wireless link is that it does not always correlate with the success rate of packets reception. RSSI simply measures the received power strength regardless of the surrounding noise. For instance, a low-strength signal in total absence of noise may have a better chance to get a higher link quality than a high-strength signal in a noisy environment.

Link Quality Indication (LQI)

The use of a Link Quality Indication (LQI) is specified by IEEE802.15.4 to assess the quality of the communication link between a receiver and transmitter [IEEE802.15.4-2006, 2006]. LQI is based on signal-to-noise ratio or energy density of the signal in the frequency band used by the standard and provides average correlation values for each incoming packet over at least 8 symbol periods. As with RSSI, LQI allows users to assess the communication link considering the environmental effects on a single transmitter/receiver pair. However, LQI provides a more thorough estimate of the quality of an IEEE 802.15.4 link than RSSI since it assesses all possible frequencies in the physical layer of the transmission.

Packet Error Rate

Packet error rate (PER) is defined as the ratio of the number of packets unsuccessfully received to the total number of packets transmitted over certain period of time. In a reliable system, it is simply expected that each data packet transmitted is received correctly by the receiver. One way to measure reliability in this manner is to keep track of the number of messages sent by the transmitter and compare the number of messages successfully received at the base station. The reliability can then be expressed as the percentage of dropped packets of data over the total number of transmissions, or as a packet error rate.

3) Power Management

In wireless sensor networks, the energy source is generally limited, often being comprised of small batteries. This limitation becomes critical when hundreds or thousands of nodes are placed in a network for long-term monitoring applications. In this circumstance, it is practically impossible to change or recharge the batteries in such a large number of nodes. From discussions with vendors and users of wireless sensors, the expected lifetime of the sensors emerged as a common concern. While there is no clear consensus on the acceptable time between battery replacement, a timeframe of 5 years appears to be a generally accepted rule-of-thumb. Many vendors claim that their sensor nodes are so efficient that no maintenance will be required for 10 years, effectively limiting the lifetime of the sensor node by the shelf life of the battery itself. Typical energy content of batteries is listed in Table 1, but self-discharge can limit the amount of energy available for useful purposes.

Table 1. Typical energy capacity of batteries [Halpern and Saleem, 2005]

Cell Size	**Energy content [mA•h]**
AAA	700
AA	1500-2000
C	5000
D	9000-12000
9V	550

To achieve long battery lives, minimal energy should be consumed by going to "sleep-mode" when not taking data, transmitting data packets for very short periods, and minimizing the amount of transmitted data. Different operation modes by a radio node are available for the management of typical power consumption, shown in Table 2. In general, signal transmission is a larger source of energy consumption than data acquisition and processing, so any efforts to minimize radio transmission will help prolong battery life. The specifics of the application, however, will have a large bearing on the time needed between battery replacements. In this sense, the lifetime of the system will be affected by the frequency of data acquisition and transmission, the power levels at which the radio transmitters are set, and the network design utilized. While a mesh network offers a potential to increase communication reliability, each node in that network serves as a repeater that must consume significantly more energy relaying messages than if it were only required to send its own data.

One advantage of sensor use in buildings is that line power is often available, and therefore, sensors can transmit data wirelessly yet get their power from a wire. Likewise, line-powered repeaters can help extend the accessibility of a wireless network without the need for multi-hop transmissions between a sensor node and a base station node. Regarding alternative means of powering sensors, significant work is underway to scavenge power from vibrations, light, or temperature gradients [Roth and Brodrick, 2008]. Such systems may eliminate chemical batteries from the sensor nodes, but some type of electrical storage will likely be needed to provide sufficient power to the sensors.

Table 2. Typical power consumption for a radio node [Zhao et al., 2002]

Radio mode	Power consumption (mW)
Transmit	15
Receive	12.5
Idle	12.4
Sleep	0.016

With power management in sensor networks, there is a need for a clear metric on the energy consumption of these sensor nodes in different applications. Battery powered nodes will require an estimate of overall energy consumption over a period of time for various activities performed by sensor nodes. Energy scavenging techniques will also require thorough studies on energy conversion mechanisms to determine the amount of energy that must be collected to permit the data acquisition and radio transmissions that are needed.

4) Interoperability

Familiarity with IEEE 802.11 and WiFi for wireless networking access in homes and offices has provided confidence for the relatively rapid adoption of wireless sensing technology in buildings. Interoperability has made WiFi attractive. Users can install wireless networking cards from a range of vendors with great confidence that reliable communication with other computers can be set up. One goal of the ZigBee alliance is to bring that same level of interoperability to the wireless sensor network community [ZigBee Alliance, 2008]. By adopting the IEEE 802.15.4 standard for the physical layer of wireless communications, ZigBee has added standards that will increase interoperability in both the networking protocol and data exchange. Thus, end users can use sensors appropriate for their particular application, and can easily integrate them with confidence into monitoring and control systems. Unless users can be sure that a wireless sensor network system can be easily modified and customized using standard components, they may be hesitant to embrace the technology in their applications.

5) Ease of Use and Maintenance

Easy deployment and minimal maintenance are also key factors that potential end users seek. Most civil and building engineers who use the wireless sensors will have little expertise in the electrical engineering and radio physics that are critical to the operation of wireless sensor networks. The wireless sensor platforms must, therefore, provide easy integration into existing networks, require minimal programming effort, possess intuitive user interfaces, and relay data in a standard format that can be easily read by applications. Additionally, these sensors must be robust enough for long-term deployment with little maintenance.

6) Security

With wireless data transmission, there is recurring concern that hackers will tweak the measured data or access building automation systems to use the wireless sensor network as a tunnel into other critical information infrastructure. Evidence of this concern is found in military facilities where the use of wireless is forbidden. Other facility managers claim that their IT security offices would raise great concerns if wireless infrastructure were installed; they have chosen not to fight that battle.

Conclusions

This paper presented key areas of concern that inhibit the use of wireless sensors in a building: 1) cost, 2) reliability, 3) power management, 4) interoperability, 5) easy of use and maintenance, and 6) security. It is proposed that measurement methods be developed to help give users a clear gauge of a system's performance in these areas. Each of the concerns may require its own set of test methods to easily assess the performance of wireless sensor networks in a particular building. A research team at NIST is currently working to develop such test methods to help users obtain a clear picture of how a wireless sensor system will work in their applications. The results from the development of test methods will provide useful guidance on the effective utilization of wireless sensor networks in a wide variety of use-cases.

References

[1] Swyt, D.A. (2007). "An Assessment of the United States Measurement System: Addressing Measurement Barriers to Accelerate Innovation," NIST Special Publication 1048, National Institute of Standards and Technology, Gaithersburg, MD.
[2] Brambley, M.R., Haves, P., McDonald, S.C., Torcellini, P., Hansen, D., Homberg, D., and Roth, K.W. (2005). "Advanced Sensors and Controls for Building Applications: Market Assessment and Potential R&D Pathways," Pacific Northwest National Laboratory Report PNNL-15149, Richland, WA.
[3] Kintner-Meyer, M. and Brambley, M. (2002). "Pros & Cons of Wireless," ASHRAE Journal, 44(11), 54-61.
[4] Lynch, J.P. (2007). "An Overview of Wireless Structural Health Monitoring for Civil Structures," Philosophical Transactions Series A: Mathematical, Physical, and Engineering Sciences, 365(1851), 345-372.
[5] IEEE 802.15.4-2006 (2006). "Wireless MAC and PHY Specifications for Low Rate Wireless Personal Area Networks (WPANs)", IEEE Computer Society, New York, NY.
[6] Halpern, M. and Saleem, K. (2005). "Battery Power Issues for WSN's," Proceedings of NICTA Formal Methods Workshop, National ICT Australia, November 7-9, Sydney, Australia.
[7] Zhao, L., Zhang, W.-H., Xu, C.-N., Xu, Y.-J. and Li, X.-W. (2002). "Energy-Aware System Design for Wireless Sensor Network," Acta Automatica Sinica, Elsevier, 32(6), 892-899.
[8] Roth, K. and Brodrick, J. (2008). "Energy Harvesting for Wireless Sensors," ASHRAE Journal, 50(5), 84-90.
[9] ZigBee Alliance (2008). <http://www.zigbee.org> (November 19, 2008).

Contour Crafting Process Planning and Optimization

Jing Zhang and Behrokh Khoshnevis

Epstein Department of Industrial & Systems Engineering, University of Southern California, Los Angeles, CA 90045 jingzhan@usc.edu, khoshnev@usc.edu

Abstract

Contour Crafting is an emerging technology that uses robotics to construct free form structures by repeatedly laying down layers of material such as concrete. The Contour Crafting technology scales up the additive fabrication process from building small industrial parts to constructing buildings. Tool path planning and optimization for Contour Crafting benefit the technology by increasing the efficiency of construction of complicated structures. This research has intended to provide a systematic solution for improving the overall system efficiency and realizing the automation of the Contour Crafting technology for building custom-designed houses. An approach is presented to find the optimal tool path for the single nozzle Contour Crafting system incorporating the physical constraints of the technology and construction considerations. Several algorithms are given to find the collision-free tool path for the multiple nozzle system based on the single nozzle approach.

1. Introduction to Contour Crafting

Contour Crafting [R2] can automatically construct custom-designed structures by repeatedly laying down construction material. It is an additive fabrication technology that uses computer control to exploit the superior surface-forming capability of troweling in order to create smooth and accurate planar and free form surfaces out of extruded materials. Unlike many other automatic additive fabrication technologies such as 3D printing, SLS, SLA, FDM[R4], which can only deliver relatively small size of three-dimensional structures (normally 1 cubic foot maximum), Contour Crafting has the capability to fabricate with thick layers using various materials and without compromising surface quality. Contour Crafting scales up the additive fabrication process to mega scale construction activities (Figure 1). The goal of Contour Crafting technology is to build custom-designed houses in a short time such as a day.

Since Contour Crafting has the ability of remarkably reducing the overall cost, injury, construction waste and impact to the environment, it can be effectively used for building houses for the low income class, shelters for disaster victims or even colonies on remote areas or other planets[R6]. Contour Crafting will also impact the construction industry for its capability and flexibility in constructing intricate or innovative structures. Its ability to build free-form shapes by utilizing the side trowels reduces the difficulty and cost of construction of complex structures. The cost of a house built by Contour Crafting technology mainly depends on the materials used and on the overall machine time. Innovative or organic form structures (such as adobe) might cost the same or even less than conventional rectangular structures because they require less support material. Architects are given more design flexibility because Contour Crafting eliminates many design limitations. CC allows architects to focus on the aesthetic appearance and functionality of the structure with less concern about construction limitations.

2. Process planning and optimization in Contour Crafting

Process planning and optimization play important roles in realizing the automation of the Contour Crafting system and improving the overall system efficiency. These functions generate optimal tool path for Contour Crafting system specific to the given structure designs. Furthermore, multiple-nozzle or multiple-gantry systems may be involved in construction of larger community and multi-residence structures. In these cases specific schedule and workload will be assigned to individual nozzles or gantries for collaborative operation. Collision between nozzles should be avoided without compromising the overall constructing efficiency. This paper intends to present a systematic methodology for Contour Crafting process planning and optimization through the following steps:

1. Describe system characteristics and define tool path elements of Contour Crafting
2. Develop practical tool path planning and an optimization method for the single nozzle CC system
3. Develop practical tool path planning and optimization methods for multi-nozzle system based on the optimization method for single nozzle

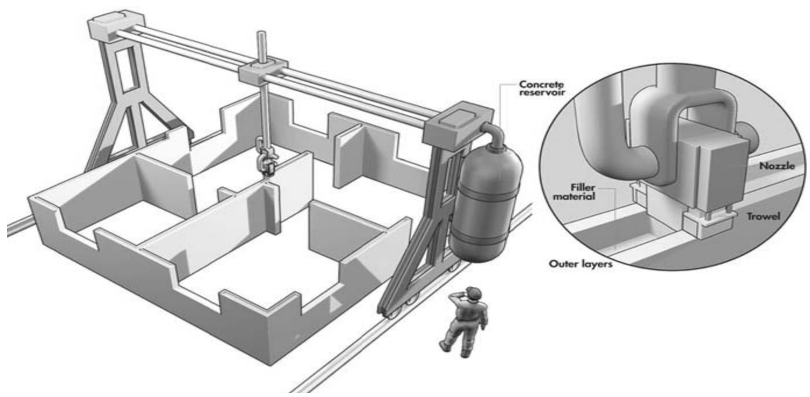


Figure 1 CC in construction operation

3. Implementation of process planning and optimization

3.1. System characteristics and tool path elements of Contour Crafting

A Contour Crafting tool path for a specific structure must describe the position, orientation, velocity, and deposition rate of the nozzle in the entire construction period. This information is then converted to a sequence of machine tasks and then fed to the Contour Crafting machine. If we define the time or energy spent on each machine task as cost in general, then the goal of optimization would be finding a path with minimum total cost associated with every machine task. Therefore, cost of deposition, airtime and other machine tasks need to be defined for calculating the overall cost for the tool path.

Cost of deposition is related to the total length of wall segments, the deposition flow rate and the moving speed of the machine. *Cost of airtime* is related to the *cost of moving* between wall segments and the *cost of rotation* along wall segments. *Cost of moving* between two segments can be determined once the distance between end points and the velocity of the machine are known. *Cost of rotation* between segments can be evaluated according to the relative orientation of the two segments. However, in the real system, the degree of rotation of the nozzle is limited because the hoses and wires attached to the nozzle may tangle and become damaged if the nozzle rotates without any limitation. For this reason a mechanical stop is used on the rotation union to prevent the nozzle from turning more than 360 degrees in either direction. Nozzle rotation direction and degree of rotation need to be adjusted if the mechanical stop impedes the re-orientation transition of the nozzle in a given direction. Therefore, cost of rotation depends on not only the rotation degree but also on the start and end positions of the stopper on the rotation union. Cost of rotation between each pair of wall segments needs to be calculated before optimization is performed.

3.2 Tool path planning and optimization method for the single nozzle CC system

Once the costs of different machine tasks and physical constraints have been defined, optimization can be performed to find the most efficient tool path for the single nozzle system. The approach presented here is to convert the CC path model to a standard TSP (traveling salesman problem). This approach considers all the possible alternatives of construction and provides optimal solution if the TSP model is solved exactly. Heuristic TSP solvers can be used if the scale of the problem is large (for instance, the structure layout has more than 1000 vertices).

TSP is used to find the shortest route to visit a collection of cities at least once and return to the starting city. In the standard TSP problem, vertices represent cities while arcs are the paths between cities. In a standard TSP, distances between two cities are the same in both directions; otherwise we have an asymmetric TSP. A solution to the TSP must return the cheapest Hamiltonian cycle of the graph which represents the cities and paths. A Hamiltonian cycle is a simple path in the graph that contains each vertex. An asymmetric TSP problem can be formulated as follows: Define $X_{ij} = 1$ (when i,j are the index of the vertices), if edge (i,j) is in the optimal tour; otherwise $X_{ij} = 0$, and $D_{ij} = d(i,j)$, when d is the traveling cost between vertices i and j. we have

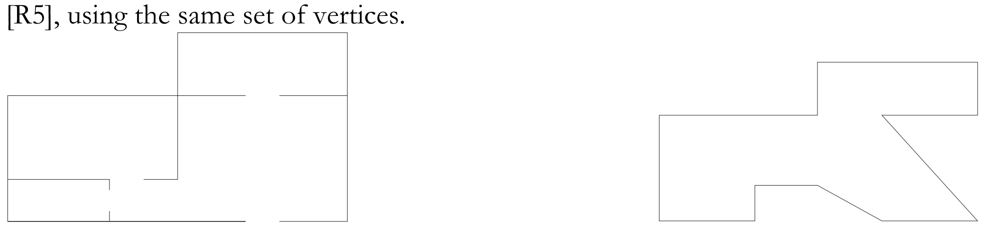
$$\text{Min } \Sigma\ \Sigma\ D_{ij} X_{ij}$$

$$\Sigma X_{ij} = 1 \qquad \text{for all } j$$

$$\Sigma X_{ij} = 1 \qquad \text{for all } I$$

$$\Sigma\Sigma X_{ij} >= 1 \qquad \text{for every } S \subseteq X \text{ (when } i \in S\ ;\ j \in X\text{-}S)$$

The graph of a structure layout cannot be directly formulated as a standard TSP problem. In the CC construction process, some edges in the graph have to be traversed by the nozzle in order to deposit concrete for building walls, which means that the CC tool path has to contain some specific edges. However, any edge can be included in the optimal path in TSP since any edge represents a path between two cities. Also, a vertex in a structure layout may have several edges incident to it, which means during the construction process, the nozzle of the CC machine will visit the same vertex more than once. However, in TSP, each vertex can be visited only once. Figure 2 shows two graphs that share the same set of vertices. One of the graphs is a structure layout for CC. Another one is the optimal TSP path generated by Concorde TSP solver [R5], using the same set of vertices.



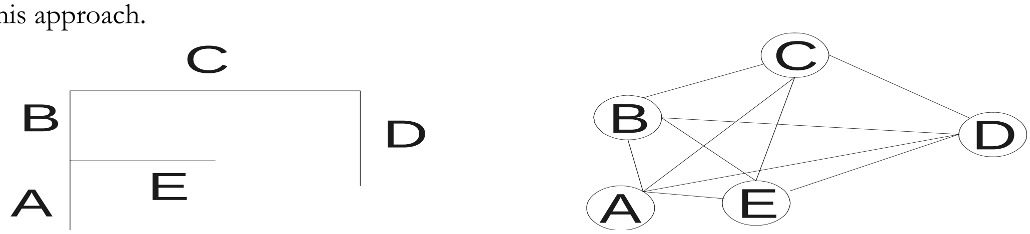
Layout model for CC construction — Tool path generated by Concorde TSP solver

Figure2: two graphs that share the same set of vertices construction

For Contour Crafting, the overall construction time of a specific structure is the sum of the overall time of concrete deposition and the overall nozzle airtime, in which the nozzle stops depositing material and travels between two deposition edges. No matter how the optimal path is generated, the nozzle should traverse all the deposition edges once and only once. The overall deposition time is determined once the structure is given. The overall nozzle idle time is the factor that determines the overall construction time for different tool paths. The optimal tool path is a path that has the minimum overall nozzle airtime. Since the nozzle of the machine can move freely in 3-dimensions, it can go straight between any vertices. The problem of finding the optimal tool path can be stated as follows:

Given a set of edges on a layout, find the optimum sequence and direction in which: (1) each edge is traversed exactly once and (2) the airtime travel (motion between two end points of two edges) is a straight line. The optimal solution minimizes the overall airtime travel.

An approach to formulate the problem is to ignore the deposition edges (walls) while only considering the traveling paths between edges (the airtime of the nozzle). In this case, walls shrink to vertices (entities), when the paths between vertices represent the cost of traveling between walls. Figure 3 shows the concept behind this approach.



A structure layout — Compact edge to vertex (letters represent edges)

Figure 3 Concept of shrinking edges

Since each edge has two vertices, the approach of shrinking the edge to a single point will have four possibilities to travel from one edge to another. As defined in Section 3.1, the cost of traveling from one wall segment to another depends on the time spent on moving and rotating the nozzle. Cost of rotation depends on the orientation of the two edges, the traveling sequence and the starting position of the stopper on the rotation union. Cost of rotation in opposite directions may be different even with the same rotation degree. Therefore, cost of rotation cannot be determined before performing the optimization. In order to formulate the problem as a TSP, some modifications need to be done:

Let Vi1 and Vi2 denote the two end points of the ith edge (i = 1; 2; . . . ; n). Let C(x , y) denote the traveling cost between points x and y, which is determined by the rotation cost and the Euclidean distance of point x and y. Define a complete network with vertex set {Vik | i = 1, 2, . . . , n; k =1,2}. Between every pair of distinct vertices (Vik , Vjl) there is an undirected edge with length given by:

C (Vik, Vjl) = - M if i = j

= Traveling cost of Vik and Vij if i ≠ j

Where M is a large number (for example, M may be set equal to the total length of any feasible tour in the original problem). For i = 1; 2; . . . ; n, the distance of -M between vertices Vi1 and Vi2 implies that the optimal tour must include the edge that connects them. Therefore, every deposition edge will be traversed by the nozzle. A minimum length Hamiltonian cycle in this network yields a practical optimal tour for the tool path optimization problem. Figure 4 and Figure 5 show the concept behind this approach

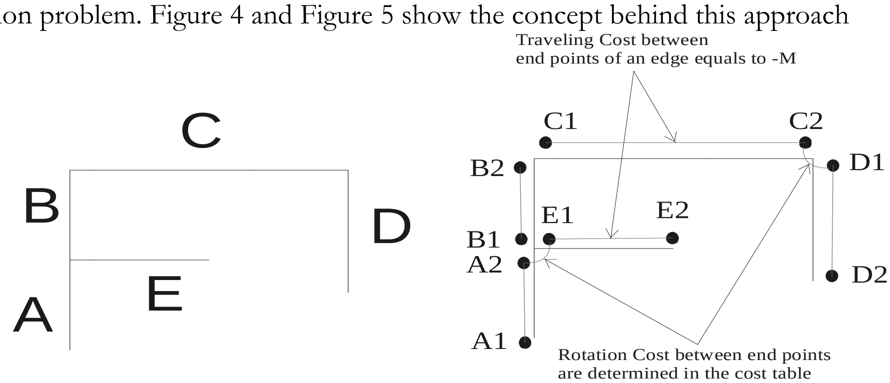


Figure 4 the concept of converting building layout to standard TSP problem.

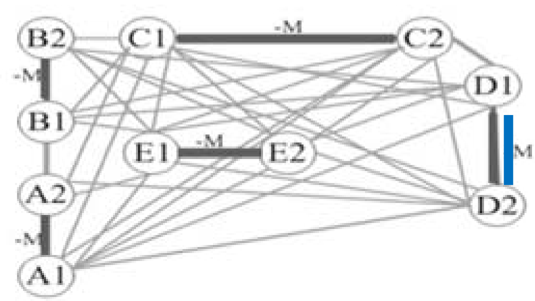


Figure 5 Traveling cost between end points of edges (blue/bold lines) equals to –M, Other traveling costs (green/thin lines) are defined in Section 3.1.

The converted CC-TSP problem can be solved by using heuristic algorithms. Most TSP solvers use effective heuristic algorithms to find the acceptable result (normally no more than 5% of the optimal solution [R3]) within reasonable time. The Lin-Kernighan algorithm [R3] has been one of the most successful tour-improving methods during the 1970's and the 1980's. The two most recent implementations of Lin-Kernighan algorithm are the Chained (sometimes also called Iterated) Lin-Kernighan algorithm by Johnson and McGeoch [R1] and the modified Lin-Kernighan algorithm introduced by Helsgaun [R1]. The former changes the classic Lin-Kernighan algorithm by having it iterating in several steps. Helsgaun make some improvements on the original Lin-Kernighan algorithm, mainly by revising restrictions and directing the search for tour parts probably belonging to the optimal solution. Helsgaun's application is used in this research.

Results: 50 structure layouts (small scale problem, less than 100 vertices) have been tested using the above approach with the given CC system parameters. Single nozzle optimal tool paths have been successfully found for all the layouts. CPLEX [R5], a commercial integer programming solver is used to check the accuracy of the result.

3.3 Tool path planning and optimization methods for multi-nozzle system based on optimization of the single nozzle case

The primary concern in using multiple nozzles (or gantries) is that collision between different nozzles/gantries should be avoided. The tool path generation of the multi-nozzle system includes two steps. The first step is to separate the original structure into different sections according to the number of nozzles by using an iterative dividing procedure. The second step is to create tool paths for these sections so that no collision between the nozzles occurs when they travel along the tool paths.

3.3.1. Step1: Iterative dividing.

In order to assign workloads to different nozzles, the original structure layout should be separated into different sections according to the number of the nozzles. Ideally, each section contains an equal amount of work load so that the construction time of all of the sections is the same. Straight lines can cut across the original layout in order to divide it into sections with the condition that the sums of the length of all of the wall segments in different sections are equal or approximate. The single nozzle optimization algorithm (CC-TSP) is applied to find out the overall construction time of each section of the layout. If the difference between the construction times is acceptable (lower than the pre-set threshold) then the workload assignment is considered to be achieved. Otherwise, the cutting lines should be moved and split the original structure, the optimization should be performed again on each section to find the difference between the construction times. The above procedures will be performed iteratively until the best result is achieved.

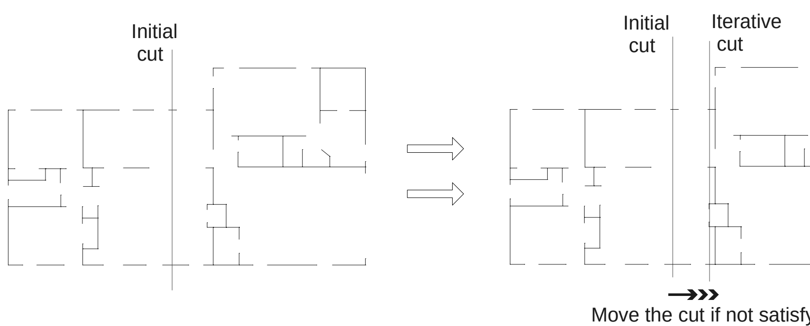


Figure 6 Iterative dividing

3.3.2. Step 2: create collision-free tool paths between the divided parts

After evenly dividing the structure into different sections, collision-free tool paths between the divided sections can be created. There are two ways to prevent collisions during the construction, they are: (1) setup a buffer area to prevent the nozzles from getting too close to each other during the construction process, and (2) analyze the x/t curves of the gantries that carry the nozzles. Three algorithms are proposed to find the optimal collision-free tool paths. Some algorithms have a higher chance of converging to a feasible solution than the others. However, the extent of optimality of their solutions might be lower. These algorithms are: (1) buffer zone; (2) path cycling; (3) buffer zone path cycling.

3.3.2.1 Buffer zone

Nozzles may collide near the shared section borders. Gantries that carry the nozzles could collide with each other when they are working near the cutting edge of adjacent sections since the width of the gantries is not equal to zero. See figure 7.

Buffer zones can be setup on both sides of the shared border in order to prevent collisions near the border. Buffer zones must meet the following conditions: (1) the size (width) of the buffer zone should be bigger than the width of the gantry; (2) the overall workload in the buffer zone should be less than half of the overall workload within the section that contains the buffer zone. When more than two gantries are working together, one gantry should avoid the collision with gantries on either side, therefore each divided section needs to have two buffer zones. The concept of auxiliary buffer zone can be used to reduce the number of buffer zones (See figure 8). First a buffer zone is generated for each section, and then the construction time of each buffer zone will be calculated. If the construction time of a buffer zone of a section is bigger than that of a buffer zone of the next section, no additional buffer zone is needed for that section. Otherwise, auxiliary buffer zones should be generated next to buffer zone with less construction time. In each section, the nozzle should work according to the following order: (1) buffer zone, (2) auxiliary buffer zone (if there is any) and (3) the main working zone. These constraints assure that the working areas of any two nozzles are mutually exclusive during the entire operation; therefore, collisions can be avoided.

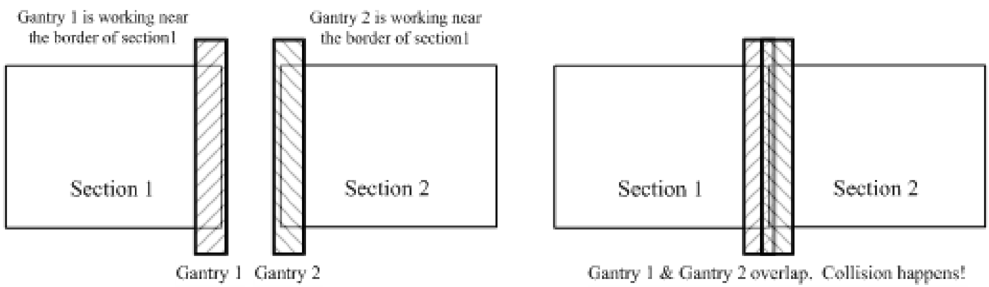


Figure 7 Possible collisions between two gantries

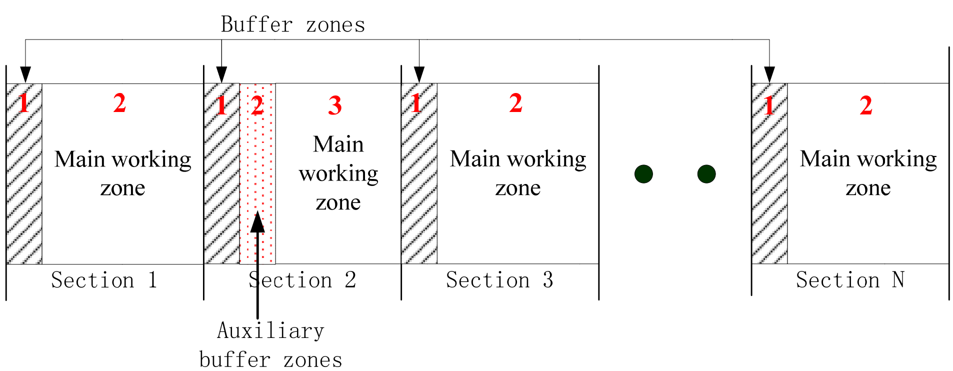


Figure 8 Auxiliary buffer zone. (numbers are the construction sequence of a nozzle)

3.3.2.2 Path cycling

Since the nozzles are carried by their corresponding gantries, two nozzles never collide if the distance between the corresponding gantries is never smaller than a specific amount in anytime. Let x1(t), x2(t) represent the x position of the nozzle 1 and nozzle 2 in time t, we have:

x1(t) – x2(t) $<$ Specific Distance (to prevent the collision), When $0<t<$ end of the construction

An x/t curve can represent the x position of a nozzle at time t. If two x/t curves never cross each other and the minimal distance between these curves is never smaller than a specific amount, then the two nozzles will not collide with each other during the entire construction process. If these two curves can be constructed and the overall time of the longer curve is minimized, then the optimal solution will be yielded. (See figure 9)

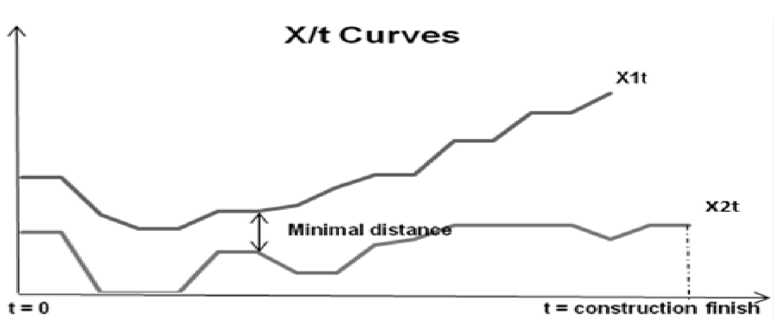


Figure 9 x/t curves

Since the optimal CC-TSP tool path is a loop (the nozzle will visit the starting point at the end), cycling the tool path will not increase the overall construction time. We can therefore always cycle one of the paths to increase the chance of finding a pair or collision-free x/t curves. To cycle a path, the first vertex is placed at the end of the sequence, but the sequence of the vertices remains the same in the tool path. The next consecutive vertex is then moved to the end of sequence if the cycling is continuous. The two paths will be checked during the cycling to see if they will collide. The cycling process will be complete if the sequence returns to its original pattern.

There are two ways to manipulate the x/t curve of the tool path to avoid collision. *Global path cycling* is achieving the result directly from the CC-TSP tool path for the single nozzle system. The optimal tool path for the entire structure is split into different tool paths according to the number of nozzles. Path cycling will be performed to find the collision-free tool path pairs between the split paths. *Individual path cycling* is similar to *global path cycling*. Instead of separating the global optimal tool path, wall segments of the original structure layout are sorted and assigned into different groups. The optimization is then performed on each of the two groups to find the local CC-TSP tool paths. These tool paths are then checked to detect if they cross each other. If they do, one of the paths is cycled and checked again until either the solution is found or the cycling process is complete (i.e., failure in finding a solution). This algorithm has a higher chance to find the result, yet the total construction time of the final solution may be longer than that of global path cycling. Figure 10 shows the concept of *individual path cycling*.

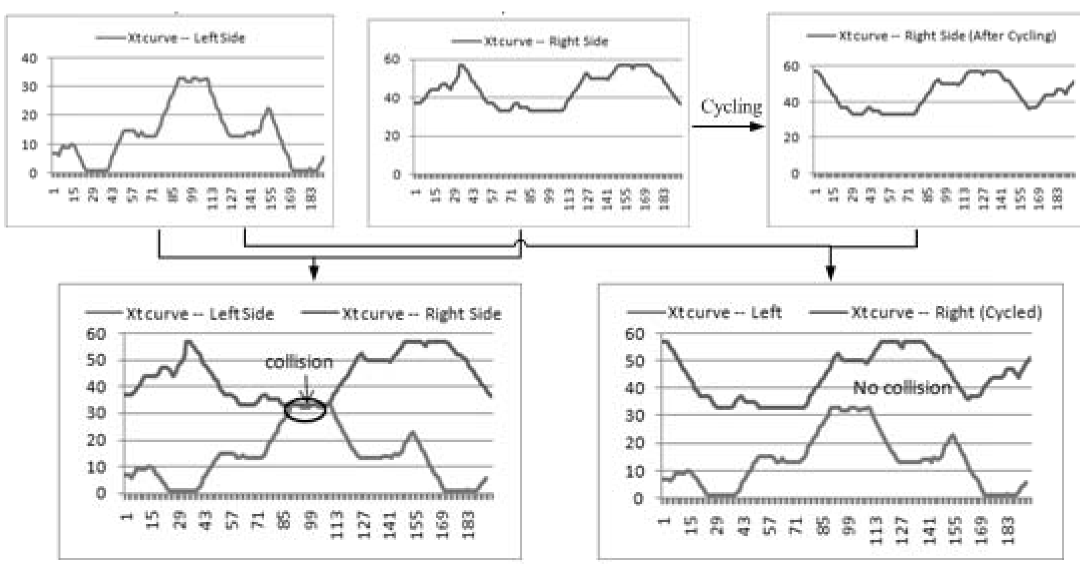


Figure 10 Individual path cycling

3.3.2.3 Buffer zone path cycling

The method of *path cycling* can create collision-free tool paths from the CC-TSP paths in most of the cases. However, the chance of finding the collision-free tool paths still depends on the geometry of the structure and the width of gantry to a certain degree. There is a higher chance to find a solution by combining the concept of the buffer zone and path cycling.

When multiple machines are used in construction, the structure is first evenly divided into sections according to the number of machines. Each section will be cut or separated into two zones: the buffer zone and the main working zone. CC-TSP tool paths will be generated for both zones of each section, and then the tool path of the working zone will be cycled to avoid collision with the buffer zone next to it. Unlike the previous method (path cycling), the procedure for making sure that the adjacent paths do not collide are independent from each other in the method of *buffer zone path cycling*. Only the paired up working zone path and buffer zone path will be checked. The purpose of cycling the path of the working zone is only to increase the chance to create collision-free tool paths with the buffer zone with which it is paired up and will not cause any collisions with any other tool paths. This will dramatically increase the chance of finding collision-free tool paths when many machines are involved in construction.

3.3.3 Results:

50 random structure layouts each having less than 200 vertices have been tested using the above algorithms with the given CC system parameters. Following are the number of collision-free tool paths for each algorithm listed:

23 for *buffer zone*, 12 for *global path cycling*, 42 for *Individual path cycling* and 50 for *buffer zone path cycling* (with the longer overall airtime than cycling based global optimization).

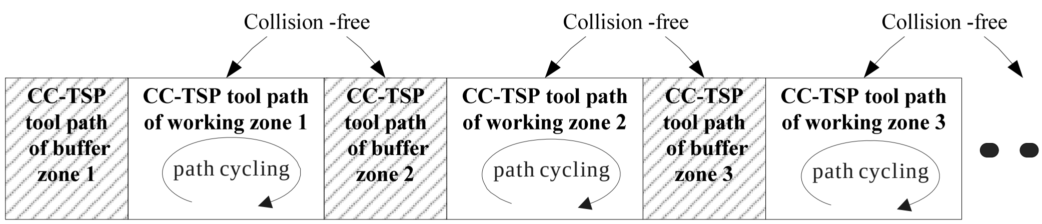


Figure 11 Buffer Zone Path Cycling for N machines system

Conclusion

This research has intended to provide a systematic solution for improving the overall efficiency of construction by Contour Crafting. An approach is presented to find the optimal tool path for the single nozzle Contour Crafting system. Several algorithms are also presented to find the collision-free tool path for the case of multiple nozzle systems. Practical and efficient tool paths can be generated using the proposed approaches to enhance the already attractive aspects of Contour Crafting.

Reference

[1] K. Helsgaun, "An effective implementation of the Lin-Kernighan traveling salesman heuristic", European Journal of Operational Research. 126, no. 1, October 2000, pp. 106-130.

[2] B. Khoshnevis, (2004) 'Automated construction by contour crafting-related robotics and information technologies' Automation in Construction, Vol. 12, pp. 5-19

[3] S. Lin, S. and B. Kernighan, "An Effective Heuristic Algorithm for the Traveling Salesman Problem", Operations Research, vol. 21, 1973, pp. 498-516.

[4] Pease III, L.F., "Rapid Prototyping Methods," ASM Handbook Volume 7:

[5] Powder Metal Technologies and Applications, 1998.

[6] Concorde TSP solver. http://www.tsp.gatech.edu/concorde.html

[7] Z. Yeh, "Trowel-Path Planning For Contour Crafting", Ph.D. Dissertation, University of Southern California, 2003

Information Strategy Planning in Construction: Framework and Processes

Boong Yeol Ryoo[1], Soo Yong Kim[2] and Young Dai Lee[3]

[1]Assistant Professor, 3137 TAMU Department of Construction Science, Texas A&M University, College Station, TX 7843-3137, U.S.A, bryoo@tamu.edu
[2]Professor, Department of Civil Engineering, Pukyong National University, San 100, Yongdangdong, Namgu Busan, 608-739, Korea, kims@pknu.ac.kr
[3]Professor, Department of Civil Engineering, Pukyong National University, San 100, Yongdangdong, Namgu Busan, 608-739, Korea, ydlee@pknu.ac.kr

Abstract

The intended objective of ISP is to develop a comprehensive plan for integrated systems development. Thus, Information Strategy Planning (ISP) is the first step in order for general contractors in the construction industry to introduce Information technology (IT) to their business and project management. Due to a lack of understanding of the role of ISP in systems development, it has been often ignored or not been considered by them. In order to design and develop an integrated system flawlessly, a well-defined software development methodology must be adopted. However, most existing methodologies are proprietary intellectual property of System Integrators (SI) whose fees are too expensive to most contractors. Their methodologies are generic and it requires a thorough understanding of their methodology and business processes in construction in order to implement in an integrated systems development in construction. This paper presents a comprehensive framework of ISP for construction contractors. Methods and tools for each step will be presented and desirable deliverables will be followed. Thus, without professional system integrators, they are able to explore information technologies that might support their business needs. It also can help them to evaluate existing systems, examine their IT and business environments, define a system architecture and functional architecture, and assess the impact on their organization.

Keywords: Information Strategy Planning (ISP), System Integration (SI), software Development Methodologies, Software Development Life Cycle (SDLC)

1. Introduction

In the construction industry, information systems have been designed and developed by computer scientists who do not have adequate understanding of the characteristics of construction business and construction projects. The construction companies also do not have sufficient knowledge of how information systems should be designed and developed in accordance with their needs. The management of system development projects must be thoroughly planned in order to put all systems and their components together [1].

Software Development Life Cycle (SDLC) models have been used in Information Technology (IT) industry. SDLC is a conceptual model for system integrators to manage software development projects. It covers from planning to operation and maintenance of a system's life cycle. However, software engineering techniques have been known to the IT industry whereas a little known to the construction industry [2]. System integrators often employ a software development methodology which has been used as a tool for producing and documenting project deliverables. Based on information engineering knowledge, comprehensive software development methodologies have been devised by system integrators. The methodologies include deliverables by phase, a project management model, a quality control model, and reasonable analysis methods and appropriate tools. Deliverables along with efficient methods and specific tools are defined and categorized throughout the software development project life cycle. Accessing such methodologies is limited since those are proprietary intellectual property of the system integrator.

Information Strategy Planning (ISP) is the basis and the first step of system integration. Without a proper ISP, system integration would not succeed [3]. Due to its characteristics, it is very hard to execute complete ISP comprehensively for construction contractors because it involves the analysis of business environment and information technology environment in the industry.

This paper presents a framework and architecture of Information Strategy Planning for general contractors. The framework is based on case studies and generic software development methodologies. It includes major tasks by phase and deliverables. Thus general contractors will be able to evaluate their deliverables.

2. Software Development Life Cycle

A framework of essential tasks to design and develop an integrated system is called "System Development Life Cycle (SDLC)". Traditionally "phased life-cycle model" and "waterfall model" have been put into practice. The models divide a software development project into six key phases such as Planning, Analysis, Design, Development, Testing, and Operation and Maintenance.

More practical models including "prototyping model," "spiral model," "iterative and incremental development model," and "component-based development" have been introduced to the information technology industry. Potential problems still exist because the models focus on design and programming and testing. Capturing industry specific business processes and knowledge has not been specified.

2.1 SDLC Methodologies

Three distinct approaches have been considered as shown in Table 1. The concept of software development methodologies has been moved forward from Process-Oriented, to Data-Oriented, and to Object-Oriented Design. In early stage of software engineering, the process modelling method was used in structured analysis and design. Yourdon/DeMarco [4] and Gane/Sarson proposed a process-driven approach to bring both data and control flow of a system together. DFD is Data Flow Diagram. James Martin [5][6] proposed a data-driven approach. Booch [7] and Yourdon/Coad [8][9] proposed objected-oriented approach.

Table 1 Characteristic of System Development Approaches

Approach	*Developers*	*Description*
Structured analysis	Yourdon/DeMarco Gane/Sarson	Unit system-oriented and Process-oriented Convert business requirements into specifications Good for computerization of existing business process
Information Engineering	James Martin	Data-oriented Plan project management throughout the project life cycle Good for enterprise system integration
Object-Oriented Analysis/Design	Grady Booch Yourdon/Coad	Object-oriented Convert functional requirements into implementation classes Good to define dependency for complex systems

2.2 CASE Tools

Even though such software development methodologies have been introduced in the information technology industry many years ago, implementing such a methodology in real projects is very complicated because of a lack of understanding of the methodologies and CASE Tools. It is also hard to measure the performance and productivity of a software development project. In many cases, success of a software development project has been relying on the knowledge, experience, and skills of team members.

As shown in Table 2, well-known system development methodologies have been introduced by business consultation firms. These are available from the firms or their business partners. Often these methodologies are not easily accessible due to expensive fees.

Method/1 focuses on three issues: what is the business environment that the client faces, in which desired direction the client would go, and what activities must be completed in order to achieve the issues [10]. Method/1 tries to help to identify functions, analyze data, and design the architecture of intended systems.

Navigator proposes system design and development tasks in more than 250 subject areas throughout project life cycle. It describes how three distinct components of system development, i.e., data, process, and information technology, can be formulated. Application Development Workbench (ADW) is a Computer-Aided Software Engineering (CASE) tool by Sterling Software and system developers are able to share resources in design and development.

In 4Front, business strategy, people, organization, processes, information technology come together. Thus understanding business strategy is important before developing a system. Information Engineering Methodology (IEM) has been adopted by Texas Instruments (TI) and it uses Information Engineering Facility (IEF) which is the first computerized methodology.

Table 2 Commercial System Development Methodologies

Methodologies	*CASE Tool*	*Developer*	*Phases*
Method/1	Foundation	Accenture	Oldest methodology and used by more than 300 organizations
Navigator	ADW	Ernest & Young	More than 250 subject areas
4Front	Excelerator	Deloitte	Planning, design and development
IEM (Information Engineering Methodology)	IEF (Information Engineering Facility)	Texas Instruments (James Martin)	Covers ISP phase to implementation phase
SSADM (LSDM)	System Engineer	LBMS	Used in the analysis and design stages of systems development
SILC (Systems Integration Life Cycle)	SHL Transform	SHL	Includes deliverables from design and development phases

3. Information Strategy Planning

Information Strategy Planning (ISP) is the process of exploring essential tasks for systems development by studying client's business plans and goals [11]. To identify major tasks and required deliverables from ISP for construction contractors, two surveys were conducted to investigate the business and information technology strategies and future plan for system development. Ninety-nine small and medium-size contractors participated in the first survey and twenty-three contractors were participated the second survey.

3.1 ISP Survey

Two surveys were sent out to small and medium size contractors and their annual volume of work is under $1.5 Billion. As shown in Figure 1, the volume of the majority of the respondents was less than $250 million which are considered as small-size forms. 31% of them were greater than $250 million but less than $750 million where as 21% of them were more than $1.5 billion.

According the survey, most contractors were not aware of such methodologies. Because of additional expenses in hiring well-known consultation firms as shown in Table 2, the contractors have been attempted to build up an information strategy planning without external assistance. This enforced them to purchase off the shelf systems for various management as shown in Table 3. This shows system integration between related systems is needed but it has not been seriously considered. One of the common misunderstanding, 57% of the respondents, was a project management system, an estimating application or a contract management application has been considered as an Enterprise Resource Planning (ERP) system. Only 14% of the respondents have been used a commercial ERP system.

For example, ERP is a key system in corporate level management. Important decisions to buy an ERP system have been made by top management (45%) and Project Management System (PMS) by the same top management (46%). Finance, accounting, human resources, and asset management had more priority than other management modules. Majority of contractors (53%) want to purchase commercial ERP systems where as only 19% of the contractors want to have custom-build management information systems. Approximately 80% of the respondents set aside a budget of less than $500,000 for an ERP system implementation even though they understand that ERP is compatible with the way of the business practice in the construction industry.

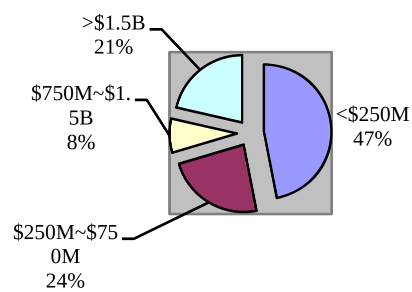


Figure 1 Size of Respondents

Table 3 Plan for System Integration

Future Plan	No. of Respondants	Percent (%)
Develop MIS	19	18%
Buy off-the-shelf ERP	53	51%
Integrate multiple modules	11	11%
Buy functional systems	16	16%
Other	4	4%

The average satisfaction rate on corporate level systems was 2.65 on 5 point scale when their integration rate was 3.28 on 5 point scale. This showed that the level of integration among the systems was poor because they did not have proper assistance from information technology professionals. The survey also showed most of contractors did not aware of the importance of information strategy planning and did not hire a standard methodology. Notably, they were not able to financially execute such tasks because their budget was limited or none.

3.2 Case Studies

The objective of this study was to identify essential tasks. Through the study a set of tasks or ISP was presented as shown in Table 3. Six construction firms were studied. Three of them were construction management firms and the other three are program management firms. One of the chief roles of program management firms is to execute heavy construction projects and development projects. Broader business environment both in national and global should be studied in their ISP. Construction management firms required the study of narrow business environment including the market analysis. Thus combining broader and barrow business environment can be beneficial for contractors to predict the image of future systems and required information technology.

3.3 Framework of ISP

In initiation phase, five phases can be considered in order to complete an information strategy planning project: Project Planning, Business Environment Analysis of Current and Future Market, Information Systems and Information Technology of Current and Future Market, Recommendations for Future System Integration, and Project Management. The project plan includes the scope of work, project team, project schedule, and the project execution plan. In order for successful project completion, key items for the project plan were considered such as project goals, project strategy and its structure, methods, staff and skills.

System integrators' perspective, the management of deliverables is critical. All tasks performed during the project must be documented. Business processes are broken down into three levels: main, group, and unit. Business actions at corporate level processes fall into main processes. These main processes are further broken into group processes for team level actions. Unit processes are to define specific procedures for individual members of the group.

In the beginning of ISP phase, the required corporate level information will be asked to complete an information strategy planning for general contractors. The defined vision and goals were among the items. Long-term and short-term plan related to the goals were collected. Especially corporate level strategy, business level strategy and organization level strategy on key business areas were defined. Corporate level strategy is to cope with where the industry is going. Business level strategy is more specific strategy and may vary depending on the types of business areas. Organization strategy is for home office management, human resources management, customer relationship management, and supply chain management. Thus, corporate level strategy should be supported by business level strategy and its functional strategy. Some certain techniques can be used for specific tasks such as Strengths, Weaknesses, Opportunities, and Threats (SWOT) analysis to evaluate the contractor's vision and goals, and plans. For public firms, it is important to validate if its vision and goals are incorporated with the national and industry trends. Through this step, some Critical Success Factor (CSF) can be identified.

4. Conclusions and Suggestions

Construction firms need a practical template so that where they can use it for system integration projects. This research presents a template that small and medium size contractors can use prior to their system integration projects. It will help the contractors to understand its current operational performance and clarify its future strategy in a systematic way.

Understand the business strategy is critical. Business environment analysis such as the identification of factors for business, IT and organization is also essential. Basic statistics tools were used to study the past trend and forecast the future trend of external business environment. Some specific methods such as SWOT, CSF, and matrix analysis were used to measure internal business environment such as business, organization and information technology.

Gap analysis was seriously considered in order for successful implementation. By benchmarking and case studies, potential mistakes can be eliminated. It helped contractors to realize the gap between desired systems and actual systems they have. Feasibility study, both technical and functional, was performed in order to prepare an IT implementation plan. Selecting a proper SDLC methodology and CASE tool is crucial even for ISP. Thus contractors can carry out ISP in a systematic way.

Acknowledgement

The authors wish to express their appreciation to all those who participate in the surveys for this research.

References

[1] Lederer, Albert L. and Sethi, V. (1996). "Key Prescriptions for Strategic Information Systems Planning," Journal of Management Information Systems, pp 35 - 62, Volume 13 , Issue 1 (June 1996).

[2] Artuk, U, Ryoo, B.Y., and Baier, R. (2007). "An Overview of Enterprise Resource Planning (ERP) Systems in Small and Medium sized Organizations: Case of Florida Contractors." Fourth International Conference on Construction in the 21st Century (CITC-IV), July 11-13, Gold Coast, Australia, 2007.

[3] Haapio, T. (2007). "A Framework for Improving Effort Management in Software Projects," Soft Process Improvement and Practice, pp 549-558, Volume 12, 2009.

[4] Zhao, S. Sun, Y, Liu, W. And Yin, M. (2007). "Research on a Framework of Information System Strategy Planning under Supply Chain Environment," International Conference on Wireless Communications, Networking and Mobile Computing, WiCom 2007, pp6266-6269, 2007.

[5] Yourdon, E. (1986) "Whatever happened to structured Nalsys," Datamation, V32, n 11 p133(4), June 1, 1986,

[6] Martin, J., Odell, J. (1992) Oriented Analysis and Design, 1992.

[7] Martin, J., Odell, J. (1993) Principles of Object-Oriented Analysis and Design, 1993.

[8] Booch, G. (1994) Object-Oriented Analysis and Design with Applications, 1994.

[9] Coad, P., and Yourdon, E. (1991) Oriented Analysis (2nd edition), 1991a.

[10] Coad, P., and Yourdon, E. (1991) Object-Oriented Design, 1991b.

[11] Andersen Consulting (1987), Foundation-Method/I: Information Planning, Version 8.0. Chicago: Andersen Consulting, 1987.

Table 4 Tasks for Information Strategy Planning by Phases

Phase	Tasks	Deliverables
Project planning	Set scope of work Organize project team and assign responsibility Call for pre-project meeting	Project plan - Scope of work (objectives), Schedule, and project team and responsibility
Project execution plan	Identify line-items Perform benefit/cost analysis, return on investment	Project execution plan - deliverables
Review of current strategy and policies	Make out corporate vision, objectives, goals, and long-term/short-term plans Investigate industry and major markets Review existing business plan Identify critical success factors	Business environment review Business strategy review Critical success factors System development plan
Review current business processes	Assess organization: functions & business processes Study industry Perform interviews Perform past trend analysis Perform SWOT Analysis Review profit models	Owner's requirements Existing business process model Future business process model
Review current information technology	Examine information management team Review applications/systems (hardware, software, and network) Evaluate information technology infrastructure Suggest suitable information technology	Report on applications Report on information technology infrastructure Analysis report on information management team
Future business processes	Study competitors, clients, and partners Identify future markets Study information technology trends Perform benchmarking Prioritize suggestions Redesign current business processes	Benchmarking report Information technology trends report Priority report Scope of work report
Future information technology	Describe information management structure Propose framework for system integration Propose information technology structure Suggest information management team	Information strategy plan System integration plan System management plan
Project Closeout	Review and submit deliverables	Documentation